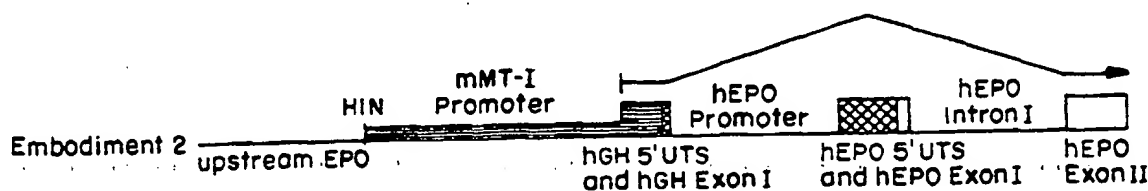




INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification⁵: C12N 15/90, A61K 48/00, C12N 15/18, 15/27	A3	(11) International Publication Number: WO 94/12650 (43) International Publication Date: 9 June 1994 (09.06.94)
(21) International Application Number: PCT/US93/11704 (22) International Filing Date: 2 December 1993 (02.12.93) (30) Priority Data: 07/985,586 3 December 1992 (03.12.92) US (71) Applicant: TRANSKARYOTIC THERAPIES, INC. [US/US]; 195 Albany Street, Cambridge, MA 02139 (US). (72) Inventors: TRECO, Douglas, A.; 87 Brantwood Road, Ar- lington, MA 02174 (US). HEARTLEIN, Michael, W.; 167 Reed Farm Road, Boxborough, MA 01719 (US). SELDEN, Richard, F.; 106 Bristol Road, Wellesley, MA 02181 (US). (74) Agents: GRANAHAN, Patricia et al.; Hamilton, Brook, Smith & Reynolds, Two Militia Drive, Lexington, MA 02173 (US).	(81) Designated States: AU, CA, JP, KR, NZ, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the</i> <i>claims and to be republished in the event of the receipt of</i> <i>amendments.</i> (88) Date of publication of the international search report: 4 August 1994 (04.08.94)	

(54) Title: ACTIVATING EXPRESSION OF AN AMPLIFYING ENDOGENOUS GENE BY HOMOLOGOUS RECOMBINATION

**(57) Abstract**

Method of activating expression of and amplifying an endogenous gene in genomic DNA of a vertebrate cell which is not expressed in the cell as obtained or is not expressed at significant levels in the cell as obtained, comprising the steps of: a) transfecting cells with DNA sequences comprising exogenous DNA sequences which repair, alter, delete or replace a sequence present in the cell or which are regulatory sequences not normally functionally linked to the endogenous gene in the cell as obtained; 2) DNA sequences homologous with genomic DNA sequences at a preselected site in the cells; and 3) amplifiable DNA encoding a selectable marker, b) maintaining the cells under conditions appropriate for homologous recombination to occur between DNA sequences homologous with genomic DNA sequences and genomic DNA sequences, c) culturing the homologously recombinant cells produced in (b) under conditions which select for amplification of the amplifiable DNA encoding a selectable marker, whereby the amplifiable DNA encoding a selectable marker and the endogenous gene functionally linked to exogenous DNA are coamplified.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

A. CLASSIFICATION OF SUBJECT MATTER
IPC 5 C12N15/90 A61K48/00 C12N15/18 C12N15/27

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 5 C12N A61K C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO,A,91 09955 (APPLIED RESEARCH SYSTEMS HOLDINGS MV) 11 July 1991	1-4, 6, 7, 9-20, 22, 29, 32, 33 1-33
Y	see the whole document ---	
Y	WO,A,90 12025 (THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY) 18 October 1990 see the whole document ---	1-33
Y	WO,A,91 00361 (THE UNITED STATES OF AMERICA DPARTEMENT OF COMMERCE) 10 January 1991 see the whole document ---	1-33
Y	US,A,4 822 736 (KELLEMS. R.E. & YEUNG, C.Y.) 18 April 1989 see the whole document ---	1-33
	--- -/--	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

3 May 1994

Date of mailing of the international search report

09-06-1994

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,
Fax (+ 31-70) 340-3016

Authorized officer

Chambonnet, F

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO,A,92 10561 (UNIVERSITY OF FLORIDA) 25 June 1992 see page 18, line 26 - line 36; claims ---	1-33
Y	WO,A,92 20808 (CELL GENESYS) 26 November 1992 see the whole document ---	1-33
Y	EP,A,0 117 059 (GENENTECH) 29 August 1984 see the whole document -----	1-33

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US93/11704

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Remark : Claims 1-14, 16-21, 26-28 and 32, 33, as far as they are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

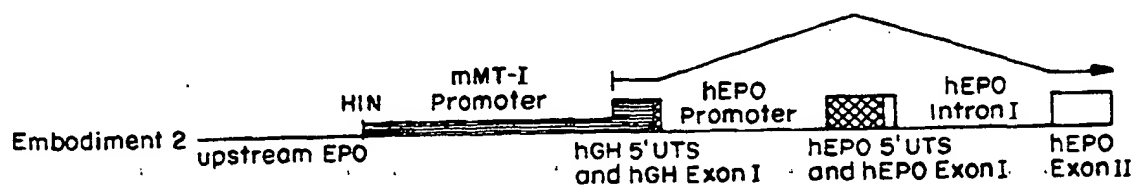
Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A-9109955	11-07-91	AU-B- 645294	13-01-94
		AU-A- 7183691	24-07-91
		EP-A- 0505500	30-09-92
		JP-T- 5504682	22-07-93
		OA-A- 9594	30-04-93
		US-A- 5272071	21-12-93
WO-A-9012025	18-10-90	US-A- 5043270	27-08-91
WO-A-9100361	10-01-91	AU-B- 636630	06-05-93
		AU-A- 5960090	17-01-91
		CA-A- 2058971	24-12-90
		EP-A- 0478696	08-04-92
		JP-T- 4502411	07-05-92
US-A-4822736	18-04-89	NONE	
WO-A-9210561	25-06-92	AU-A- 9098891	08-07-92
		EP-A- 0560885	22-09-93
WO-A-9220808	26-11-92	AU-A- 2140292	30-12-92
		CA-A- 2092695	16-11-92
		EP-A- 0539573	05-05-93
EP-A-0117059	29-08-84	AU-B- 572108	05-05-88
		AU-A- 2353484	26-07-84
		CA-A- 1271149	03-07-90
		JP-A- 59183693	18-10-84
		US-A- 5011795	30-04-91
		US-A- 5010002	23-04-91
		US-A- 5268291	07-12-93



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁵: C12N 15/90, A61K 48/00, C12N 15/18, 15/27	A2	(11) International Publication Number: WO 94/12650 (43) International Publication Date: 9 June 1994 (09.06.94)
(21) International Application Number: PCT/US93/11704 (22) International Filing Date: 2 December 1993 (02.12.93) (30) Priority Data: 07/985,586 3 December 1992 (03.12.92) US (71) Applicant: TRANSKARYOTIC THERAPIES, INC. [US/US]; 195 Albany Street, Cambridge, MA 02139 (US). (72) Inventors: TRECO, Douglas, A.; 87 Brantwood Road, Ar- lington, MA 02174 (US). HEARTLEIN, Michael, W.; 167 Reed Farm Road, Boxborough, MA 01719 (US). SELDEN, Richard, F.; 106 Bristol Road, Wellesley, MA 02181 (US). (74) Agents: GRANAHAN, Patricia et al.; Hamilton, Brook, Smith & Reynolds, Two Militia Drive, Lexington, MA 02173 (US).		(81) Designated States: AU, CA, JP, KR, NZ, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>Without international search report and to be republished</i> <i>upon receipt of that report.</i>

(54) Title: **ACTIVATING EXPRESSION OF AN AMPLIFYING ENDOGENOUS GENE BY HOMOLOGOUS RECOMBINATION**

**(57) Abstract**

Method of activating expression of and amplifying an endogenous gene in genomic DNA of a vertebrate cell which is not expressed in the cell as obtained or is not expressed at significant levels in the cell as obtained, comprising the steps of: (a) transfecting cells with DNA sequences comprising: 1) exogenous DNA sequences which repair, alter, delete or replace a sequence present in the cell or which are regulatory sequences not normally functionally linked to the endogenous gene in the cell as obtained; 2) DNA sequences homologous with genomic DNA sequences at a preselected site in the cells; and 3) amplifiable DNA encoding a selectable marker; (b) maintaining the cells under conditions appropriate for homologous recombination to occur between DNA sequences homologous with genomic DNA sequences and genomic DNA sequences; (c) culturing the homologously recombinant cells produced in (b) under conditions which select from amplification of the amplifiable DNA encoding a selectable marker, whereby the amplifiable DNA encoding a selectable marker and the endogenous gene functionally linked to exogenous DNA of (a) 1) are coamplified.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

ACTIVATING EXPRESSION OF AN AMPLIFYING ENDOGENOUS GENE BY HOMOLOGOUS RECOMBINATION.

Background of the Invention

Current approaches to treating disease by administering therapeutic proteins include in vitro production of therapeutic proteins for conventional pharmaceutical delivery (e.g. intravenous, subcutaneous, or intramuscular injection) and, more recently, gene therapy.

Proteins of therapeutic interest are generally produced by introducing exogenous DNA encoding the protein of therapeutic interest into appropriate cells. Presently-available approaches to gene therapy make use of infectious vectors, such as retroviral vectors, which include the genetic material to be expressed. Such approaches have limitations, such as the potential of generating replication-competent virus during vector production; recombination between the therapeutic virus and endogenous retroviral genomes, potentially generating infectious agents with novel cell specificities, host ranges, or increased virulence and cytotoxicity; independent integration into large numbers of cells, increasing the risk of a tumorigenic insertional event; limited cloning capacity in the retrovirus (which restricts therapeutic applicability) and short-lived in vivo expression of the product of interest. A better approach to providing gene products, particularly one which avoids the risks associated with presently available methods and provides long-term treatment, would be valuable.

Summary of the Invention

The present invention relates to improved methods for both the in vitro production of therapeutic proteins and for the production and delivery of therapeutic proteins by

-2-

gene therapy. The present method describes an approach which activates expression of endogenous cellular genes, and further allows amplification of the activated endogenous cellular genes, which does not require in vitro manipulation and transfection of exogenous DNA encoding proteins of therapeutic interest.

The present invention relates to transfected cells, both transfected primary or secondary cells (i.e., non-immortalized cells) and transfected immortalized cells, useful for producing proteins, particularly therapeutic proteins, methods of making such cells, methods of using the cells for in vitro protein production and methods of gene therapy. Cells of the present invention are of vertebrate origin, particularly of mammalian origin and even more particularly of human origin. Cells produced by the method of the present invention contain exogenous DNA which encodes a therapeutic product, exogenous DNA which is itself a therapeutic product and/or exogenous DNA which causes the transfected cells to express a gene at a higher level or with a pattern of regulation or induction that is different than occurs in the corresponding nontransfected cell.

The present invention also relates to methods by which primary, secondary, and immortalized cells are transfected to include exogenous genetic material, methods of producing clonal cell strains or heterogenous cell strains, and methods of immunizing animals, or producing antibodies in immunized animals, using the transfected primary, secondary, or immortalized cells.

The present invention relates particularly to a method of gene targeting or homologous recombination in cells of vertebrate, particularly mammalian, origin. That is, it relates to a method of introducing DNA into primary, secondary, or immortalized cells of vertebrate origin

-3-

through homologous recombination, such that the DNA is introduced into genomic DNA of the primary, secondary, or immortalized cells at a preselected site. The targeting sequences used are determined by (selected with reference to) the site into which the exogenous DNA is to be inserted. The present invention further relates to homologously recombinant primary, secondary, or immortalized cells, referred to as homologously recombinant (HR) primary, secondary or immortalized cells, produced by the present method and to uses of the HR primary, secondary, or immortalized cells.

The present invention also relates to a method of activating (i.e., turning on) a gene present in primary, secondary, or immortalized cells of vertebrate origin, which is normally not expressed in the cells or is not expressed at physiologically significant levels in the cells as obtained. According to the present method, homologous recombination is used to replace or disable the regulatory region normally associated with the gene in cells as obtained with a regulatory sequence which causes the gene to be expressed at levels higher than evident in the corresponding nontransfected cell, or to display a pattern of regulation or induction that is different than evident in the corresponding nontransfected cell. The present invention, therefore, relates to a method of making proteins by turning on or activating an endogenous gene which encodes the desired product in transfected primary, secondary, or immortalized cells.

In one embodiment, the activated gene can be further amplified by the inclusion of a selectable marker gene which has the property that cells containing amplified copies of the selectable marker gene can be selected for by culturing the cells in the presence of the appropriate selectable agent. The activated endogenous gene which is

-4-

near or linked to the amplified selectable marker gene will also be amplified in cells containing the amplified selectable marker gene. Cells containing many copies of the activated endogenous gene are useful for in vitro protein production and gene therapy.

Gene targeting and amplification as disclosed in the present invention are particularly useful for turning on the expression of genes which form transcription units which are sufficiently large that they are difficult to isolate and express, or for turning on genes for which the entire protein coding region is unavailable or has not been cloned. The present invention also describes a method by which homologous recombination is used to convert a gene into a cDNA copy, devoid of introns, for transfer into yeast or bacteria for in vitro protein production.

Transfected cells of the present invention are useful in a number of applications in humans and animals. In one embodiment, the cells can be implanted into a human or an animal for protein delivery in the human or animal. For example, human growth hormone (hGH), human EPO (hEPO), human insulinotropin and other proteins can be delivered systemically or locally in humans for therapeutic benefits. Barrier devices, which contain transfected cells which express a therapeutic product and through which the therapeutic product is freely permeable, can be used to retain cells in a fixed position in vivo or to protect and isolate the cells from the host's immune system. Barrier devices are particularly useful and allow transfected immortalized cells, transfected cells from another species (transfected xenogeneic cells), or cells from a nonhistocompatibility-matched donor (transfected allogeneic cells) to be implanted for treatment of human or animal conditions or for agricultural uses (e.g., meat and dairy

-5-

production). Barrier devices also allow convenient short-term (i.e., transient) therapy by providing ready access to the cells for removal when the treatment regimen is to be halted for any reason. Transfected xenogeneic and
5 allogeneic cells may be used for short-term gene therapy, such that the gene product produced by the cells will be delivered in vivo until the cells are rejected by the host's immune system.

Transfected cells of the present invention are also
10 useful for eliciting antibody production or for immunizing humans and animals against pathogenic agents. Implanted transfected cells can be used to deliver immunizing antigens that result in stimulation of the host's cellular and humoral immune responses. These immune responses can be
15 designed for protection of the host from future infectious agents (i.e., for vaccination), to stimulate and augment the disease-fighting capabilities directed against an ongoing infection, or to produce antibodies directed
20 against the antigen produced in vivo by the transfected cells that can be useful for therapeutic or diagnostic purposes. Removable barrier devices can be used to allow a simple means of terminating exposure to the antigen. Alternatively, the use of cells that will ultimately be rejected (xenogeneic or allogeneic transfected cells) can
25 be used to limit exposure to the antigen, since antigen production will cease when the cells have been rejected.

The methods of the present invention can be used to produce primary, secondary, or immortalized cells producing a wide variety of therapeutically useful products,
30 including (but not limited to): hormones, cytokines, antigens, antibodies, enzymes, clotting factors, transport proteins, receptors, regulatory proteins, structural proteins, transcription factors, or anti-sense RNA. Additionally, the methods of the present invention can be

-6-

used to produce cells which produce non-naturally occurring ribozymes, proteins, or nucleic acids which are useful for in vitro production of a therapeutic product or for gene therapy.

5 Brief Description of the Drawings

Figure 1 is a schematic representation of plasmid pXGH5, which includes the human growth hormone (hGH) gene under the control of the mouse metallothionein promoter.

10 Figure 2 is a schematic representation of plasmid pCDNEO, which includes the neo coding region (BamHI-BglII fragment) from plasmid pSV2neo inserted into the BamHI site of plasmid pCD; the Amp-R and pBR322ori sequences from pBR322; and the polyA, 16S splice junctions and early promoter regions from SV40.

15 Figure 3 is a schematic representation of plasmid pXEP01. The solid black arc represents the pUC12 backbone and the arrow denotes the direction of transcription of the ampicillin resistance gene. The stippled arc represents the mouse metallothionein promoter (pmMT1). The
20 unfilled arc interrupted by black boxes represents the human erythropoietin EPO gene (the black boxes denote exons and the arrow indicates the direction hEPO transcription). The relative positions of restriction endonuclease recognition sites are indicated.

25 Figure 4 is a schematic representation of plasmid pE3neoEPO. The positions of the human erythropoietin gene and the neo and amp resistance genes are indicated. Arrows indicate the directions of transcription of the various genes. pmMT1 denotes the mouse metallothionein
30 promoter (driving hEPO expression) and pTK denotes the Herpes Simplex Virus thymidine kinase promoter (driving neo expression). The dotted regions of the map mark the positions of human HPRT sequences. The relative positions

-7-

of restriction endonuclease recognition sites are indicated.

Figure 5 is a schematic diagram of a strategy for transcriptionally activating the hEPO gene; the thin lines represent hEPO sequences; thick lines, mouse metallothionein I promoter; stippled box, 5' untranslated region of hGH; solid box, hGH exon 1; open boxes, hEPO coding sequences; HIII, HindIII site.

Figure 6 is a schematic diagram of a strategy for transcriptionally activating the hEPO gene; the thin lines represent hEPO sequences; thick lines, mouse metallothionein I promoter; stippled box, 5' untranslated region of hGH; solid box, hGH exon 1; striped box, 10 bp linker from hEPO intron 1; cross-hatched box, 5' untranslated region of hEPO; and open boxes, hEPO coding sequences; HIII, HindIII site.

Figure 7 is a graphic representation of erythropoietin expression in a targeted human cell line subjected to stepwise selection in 0.02, 0.05, 0.1, 0.2 and 0.4 μ M methotrexate.

Figure 8 is a schematic representation of plasmid pREPO4.

Detailed Description of the Invention

Overview of the Invention

The present invention and the methods described in the applications incorporated herein by reference relate to transfected primary, secondary, and immortalized cells of vertebrate origin, particularly mammalian origin, transfected with exogenous genetic material (DNA or RNA) which encodes a clinically useful product, methods by which primary, secondary, and immortalized cells are transfected to include exogenous genetic material, methods of producing clonal cell strains or heterogenous cell

-8-

strains which express exogenous genetic material, a method of providing clinically useful products in physiologically useful quantities to an individual in need thereof through the use of transfected cells of the present invention, methods of vaccinating animals for protection against pathogenic viruses or microbial agents expressing epitopes antigenically related to products expressed by the transfected cells and methods of producing antibodies directed against a product made by the transfected primary, secondary, or immortalized cells. Clinically useful products can be produced in vitro, by purification from the transfected cells, or produced in vivo, by implantation into a non-human animal or human (i.e., gene therapy). Whether produced in vitro or in vivo, the clinically useful products can include hormones, cytokines, antigens, antibodies, enzymes, clotting factors, transport proteins, receptors, regulatory proteins, structural proteins, transcription factors, anti-sense RNA. Additionally, the methods of the present invention can be used to produce cells which produce non-naturally occurring ribozymes, proteins, or nucleic acids.

In one embodiment, the present invention relates to a method of gene or DNA targeting in cells of vertebrate, particularly mammalian, origin. That is, it relates to a method of introducing DNA into primary, secondary, or immortalized cells of vertebrate origin through homologous recombination or targeting of the DNA, which is introduced into genomic DNA of the primary, secondary, or immortalized cells at a preselected site. The targeting sequences used are determined by (selected with reference to) the site into which the exogenous DNA is to be inserted. The present invention further relates to homologously recombinant primary, secondary or immortalized cells, referred to as homologously recombinant (HR) primary, secondary or

-9-

immortalized cells, produced by the present method and to uses of the HR primary, secondary, or immortalized cells.

The present invention also relates to a method of activating a gene which is present in primary cells,
5 secondary cells or immortalized cells of vertebrate origin, but is normally not expressed in the cells or is not expressed at significant levels in the cells. Homologous recombination or targeting is used to replace or disable the regulatory region normally associated with the gene
10 with a regulatory sequence which causes the gene to be expressed at levels higher than evident in the corresponding nontransfected cell, or causes the gene to display a pattern of regulation or induction that is different than evident in the corresponding nontransfected cell. The
15 present invention, therefore, relates to a method of making proteins by activating an endogenous gene which encodes the desired product in transfected primary, secondary or immortalized cells.

Several embodiments in which exogenous DNA undergoes
20 homologous recombination with genomic DNA of transfected (recipient) cells can be practiced according to the present invention. In one embodiment, introduction of the exogenous DNA results in the activation of a gene that is normally not expressed or expressed in levels too low to
25 be useful for in vitro protein production or gene therapy. In a second embodiment, sequences encoding a product of therapeutic utility are directed to integrate into the recipient cell genome via homologous recombination at a preselected site in the recipient cell's genome, such that
30 the site of integration is precisely known and the site can be chosen for its favorable properties (e.g., the site allows for high levels of expression of exogenous DNA).

In a third embodiment, the present invention describes a method of activating (i.e turning on) and ampli-

-10-

5 fying an endogenous gene encoding a desired product in a
transfected, primary, secondary, or immortalized cell.
That is, it relates to a method of introducing, by homologous
recombination with genomic DNA, DNA sequences which
10 are not normally functionally linked to the endogenous
gene and (1) which, when inserted into the host genome at
or near the endogenous gene, serve to alter (e.g., activate)
the expression of the endogenous gene, and further
15 (2) allow for selection of cells in which the activated
endogenous gene is amplified. Amplifiable DNA sequences
useful in the present invention include, but are not
limited to, sequences which encode the selectable markers
dihydrofolate reductase, adenosine deaminase, and the CAD
20 gene (encoding the trifunctional protein carbamyl phosphate
synthase, aspartate transcarbamylase, and dihydroorotase).
Improved versions of these sequences and other
amplifiable sequences can also be used. According to the
present method, the amplifiable DNA sequences encoding a
selectable marker and the DNA sequences which alter the
25 regulation of expression of the endogenous gene are introduced
into the primary, secondary, or immortalized cell in
association with DNA sequences homologous to genomic DNA
sequences at a preselected site in the cell's genome.
This site will generally be within or upstream of a gene
30 encoding a therapeutic product or at a site that affects
the desired gene's function. The DNA sequences which
alter the expression of the endogenous gene, the amplifiable
sequences which encode a selectable marker, and the
sequences which are homologous to a preselected site in
30 genomic DNA can be introduced into the primary, secondary,
or immortalized cell as a single DNA construct, or as
separate DNA sequences which become physically linked in
the genome of a transfected cell. Further, the DNA can be
introduced as linear, double stranded DNA, with or without

-11-

single strand d regions at one or both ends, or the DNA can be introduced as circular DNA. After the exogenous DNA is introduced into the cell, the cell is maintained under conditions appropriate for homologous recombination to occur between the genomic DNA and a portion of the introduced DNA. Homologous recombination between the genomic DNA and the introduced DNA results in a homologically recombinant primary, secondary, or immortalized cell in which sequences which alter the expression of an endogenous gene, and the amplifiable sequences encoding a selectable marker, are operatively linked to an endogenous gene encoding a therapeutic product. Culturing the resulting homologically recombinant cell under conditions which select for amplification of the amplifiable DNA encoding a selectable marker results in a cell containing an amplified selectable marker and a coamplified endogenous gene whose expression has been altered. Cells produced by this method can be cultured under conditions suitable for the expression of the therapeutic protein, thereby producing the therapeutic protein in vitro, or the cells can be used for in vivo delivery of a therapeutic protein (i.e., gene therapy).

Additional embodiments are possible. The targeting event can be a simple insertion of a regulatory sequence, placing the endogenous gene under the control of the new regulatory sequence (for example, by insertion of either a promoter or an enhancer, or both, upstream of an endogenous gene). The targeting event can be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. The targeting event can replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally-occurring elements, or displays a pattern of regula-

-12-

tion or induction that is different from the corresponding nontransfected cell. In this embodiment the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event can be facilitated by the use of one or more selectable marker genes that are physically associated with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the host cell genome. The identification of the targeting event can also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The present invention also relates to a method by which homologous recombination is used to convert a gene into a cDNA copy (a gene copy devoid of introns). The cDNA copy can be transferred into yeast or bacteria for in vitro protein production, or the cDNA copy can be inserted into a mammalian cell for protein production. If the cDNA is to be transferred to microbial cells, two DNA constructs containing targeting sequences are introduced by homologous recombination, one construct upstream of and one construct downstream of a human gene encoding a therapeutic protein. The sequences introduced upstream include DNA sequences homologous to genomic DNA sequences at or upstream of the DNA encoding the first amino acid of a mature, processed therapeutic protein; a retroviral LTR;

-13-

sequences encoding a marker for selection in microbial cells; a regulatory element that functions in microbial cells; and DNA encoding a leader peptide that promotes secretion from microbial cells. The sequences introduced upstream are introduced near to and upstream of genomic DNA encoding the first amino acid of a mature, processed therapeutic protein. The sequences introduced downstream include DNA sequences homologous to genomic DNA sequences at or downstream of the DNA encoding the last amino acid of a mature, processed protein; a microbial transcriptional termination sequence; sequences capable of directing DNA replication in microbial cells; and a retroviral LTR. The sequences introduced downstream are introduced adjacent to and downstream of the DNA encoding the stop codon of the mature, processed therapeutic protein. After introducing into the cells each of the two DNA constructs, the cells are maintained under conditions appropriate for homologous recombination between the introduced DNA and genomic DNA, thereby producing homologously recombinant cells. Optionally, one or both of the DNA constructs can encode one or more markers for either positive or negative selection of cells containing the DNA construct, and a selection step can be added to the method after one or both of the DNA constructs have been introduced into the cells. Alternatively, the sequences encoding the marker for selection in microbial cells and the sequences capable of directing DNA replication in microbial cells can both be present in either the upstream or the downstream targeting construct, or the marker for selection in microbial cells can be present in the downstream targeting construct and the sequences capable of directing DNA replication in microbial cells can be present in the upstream targeting construct. The homologously recombinant cells are then cultured under conditions appropriate for LTR directed

-14-

transcription, processing and reverse transcription of the RNA product of the gene encoding the therapeutic protein. The product of reverse transcription is a DNA construct comprising an intronless DNA copy encoding the therapeutic protein, operatively linked to DNA sequences comprising the two exogenous DNA constructs described above. The intronless DNA construct produced by the present method is then introduced into a microbial cell. The microbial cell is then cultured under conditions appropriate for expression and secretion of the therapeutic protein.

Transfected Cells

As used herein, the term primary cell includes cells present in a suspension of cells isolated from a vertebrate tissue source (prior to their being plated, i.e., attached to a tissue culture substrate such as a dish or flask), cells present in an explant derived from tissue, both of the previous types of cells plated for the first time, and cell suspensions derived from these plated cells. The term secondary cell or cell strain refers to cells at all subsequent steps in culturing. That is, the first time a plated primary cell is removed from the culture substrate and replated (passaged), it is referred to herein as a secondary cell, as are all cells in subsequent passages. Secondary cells are cell strains which consist of secondary cells which have been passaged one or more times. A cell strain consists of secondary cells that: 1) have been passaged one or more times; 2) exhibit a finite number of mean population doublings in culture; 3) exhibit the properties of contact-inhibited, anchorage dependent growth (anchorage-dependence does not apply to cells that are propagated in suspension culture); and 4) are not immortalized.

-15-

Cells transfected by the subject method fall into four types or categories: 1) cells which do not, as obtained, make or contain the therapeutic product, 2) cells which make or contain the therapeutic product but in smaller quantities than normal (in quantities less than the physiologically normal lower level) or in defective form, 3) cells which make the therapeutic product at physiologically normal levels, but are to be augmented or enhanced in their content or production, and 4) cells in which it is desirable to change the pattern of regulation or induction of a gene encoding a therapeutic product.

Primary and secondary cells to be transfected by the present method can be obtained from a variety of tissues and include all cell types which can be maintained in culture. For example, primary and secondary cells which can be transfected by the present method include fibroblasts, keratinocytes, epithelial cells (e.g., mammary epithelial cells, intestinal epithelial cells), endothelial cells, glial cells, neural cells, formed elements of the blood (e.g., lymphocytes, bone marrow cells), muscle cells and precursors of these somatic cell types. Primary cells are preferably obtained from the individual to whom the transfected primary or secondary cells are administered. However, primary cells can be obtained from a donor (other than the recipient) of the same species or another species (e.g., mouse, rat, rabbit, cat, dog, pig, cow, bird, sheep, goat, horse).

Transfected primary and secondary cells have been produced, with or without phenotypic selection and shown to express exogenous DNA encoding a therapeutic product including, for example, hGH, EPO and insulinotropin.

-16-

Immortalized cells can also be transfected by the present method and used for either protein production or gene therapy. Examples of immortalized human cell lines useful for protein production or gene therapy by the present method include, but are not limited to, HT1080, HeLa, MCF-7 breast cancer cells, K-562 leukemia cells, KB carcinoma cells, 2780AD ovarian carcinoma cells, Raji (ATCC CCL 86) cells, Jurkat (ATCC TIB 152) cells, Namalwa (ATCC CRL 1432) cells, HL-60 (ATCC CCL 240) cells, Daudi (ATCC CCL 213) cells, RPMI 8226 (ATCC CCL 155) cells and MOLT-4 (ATCC CRL 1582) cells. Immortalized cells from other species (e.g., chinese hamster ovary (CHO) cells or mouse L cells) can be used for in vitro protein production or gene therapy. In addition, primary or secondary human cells, as well as primary or secondary cells from other species which display the properties of gene amplification in vitro can be used for in vitro protein production or gene therapy.

Exogenous DNA

Exogenous DNA incorporated into primary, secondary or immortalized cells by the present method is: 1) DNA which encodes a translation or transcription product whose expression in cells is desired, or a portion of a translation or transcription product, such as a protein product or RNA product useful to treat an existing condition or prevent it from occurring (eg., hGH, EPO or insulinotro-pin); or 2) DNA which does not encode a gene product but is itself useful, such as a transcriptional regulatory sequence or DNA useful to treat an existing condition or prevent it from occurring.

DNA transfected into primary, secondary or immortalized cells can encode an entire desired product, or can encode, for example, the active or functional portion(s)

-17-

of the product. The product can be, for example, a hormone, a cytokine, an antigen, an antibody, an enzyme, a clotting factor, a transport protein, a receptor, a regulatory protein, a structural protein, a transcription factor, an anti-sense RNA, or a ribozyme. Additionally, the product can be a protein or a nucleic acid which does not occur in nature (i.e., a novel protein or novel nucleic acid). The DNA can be obtained from a source in which it occurs in nature or can be produced, using genetic engineering techniques or synthetic processes. The DNA can encode one or more therapeutic products. After transfection, the exogenous DNA is stably incorporated into the recipient cell's genome (along with any additional sequences present in the DNA construct used), from which it is expressed or otherwise functions. Alternatively, the exogenous DNA can be used to target to DNA that exists episomally within cells.

DNA encoding the desired product can be introduced into cells under the control of an inducible promoter, with the result that cells as produced or as introduced into an individual do not express the product but can be induced to do so (i.e., production is induced after the transfected cells are produced but before implantation or after implantation). DNA encoding the desired product can, of course, be introduced into cells in such a manner that it is expressed upon introduction (i.e., without induction).

As taught herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences can be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation

-18-

sites, regulatory protein binding sites or combinations of said sequences. (Alternatively, sequences which affect the structure or stability of the RNA or protein produced can be replaced, removed, added, or otherwise modified by targeting, these sequences including polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules). According to this method, introduction of the exogenous DNA results in disablement of the endogenous sequences which control expression of the endogenous gene, either by replacing all or a portion of the endogenous (genomic) sequence or otherwise disrupting the function of the endogenous sequence. In the situation where targeting is used to replace a protein coding domain, chimeric, multifunctional proteins can be produced which combine structural, enzymatic, or ligand or receptor binding properties from two or more proteins into one polypeptide.

20 Selectable Markers

A variety of selectable markers can be incorporated into primary, secondary or immortalized cells. For example, a selectable marker which confers a selectable phenotype such as drug resistance, nutritional auxotrophy, resistance to a cytotoxic agent or expression of a surface protein, can be used. Selectable marker genes which can be used include neo, gpt, dhfr, ada, pac, hyg, CAD, and hisD. The selectable phenotype conferred makes it possible to identify and isolate recipient cells. Amplifiable genes encoding selectable markers (e.g., ada, dhfr and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydro-orotase) have the added characteristic that they enable the selec-

-19-

tion of cells containing amplified copies of the selectable marker inserted into the genome. This feature provides a mechanism for significantly increasing the copy number of an adjacent or linked gene for which amplification is desirable.

Selectable markers can be divided into two categories: positively selectable and negatively selectable (in other words, markers for either positive selection or negative selection). In positive selection, cells expressing the positively selectable marker are capable of surviving treatment with a selective agent (such as neo, gpt, dhfr, ada, pac, hyg, mdrl and hisD). In negative selection, cells expressing the negatively selectable marker are destroyed in the presence of the selective agent (e.g., tk, gpt).

DNA Constructs

DNA constructs, which include exogenous DNA and, optionally, DNA encoding a selectable marker, along with additional sequences necessary for expression of the exogenous DNA in recipient cells, are used to transfect primary, secondary or immortalized cells in which the encoded product is to be produced. The DNA construct can also include targeting sequences for homologous recombination with host cell DNA. DNA constructs which include exogenous DNA sequences which do not encode a gene product (and are the therapeutic product) and, optionally, include DNA encoding a selectable marker, can be used to transfect primary, secondary or immortalized cells. The DNA constructs may be introduced into cells by a variety of methods, including electroporation, microinjection, calcium phosphate precipitation, and liposome- polybrene- or DEAE dextran-mediated transfection. Alternatively, infectious vectors, such as retroviral, herpes, adeno-virus,

-20-

adenovirus-associated, mumps and poliovirus vectors, can be used to introduce the DNA.

In one embodiment, the DNA construct includes exogenous DNA and one or more targeting sequences, generally located at both ends of the exogenous DNA sequence. Targeting sequences are DNA sequences normally present in the genome of the cells as obtained (e.g., an essential gene, a nonessential gene or noncoding DNA, or sequences present in the genome through a previous modification). Such a construct is useful to integrate exogenous DNA (at a preselected site in a recipient cell is genome) encoding a therapeutic product, such as a hormone, a cytokine, an antigen, an antibody, an enzyme, a clotting factor, a transport protein, a receptor, a regulatory protein, a structural protein, an anti-sense RNA, a ribozyme or a protein or a nucleic acid which does not occur in nature. In particular, exogenous DNA can encode one of the following: Factor VIII, Factor IX, erythropoietin, alpha-1 antitrypsin, calcitonin, glucocerebrosidase, growth hormone, low density lipoprotein (LDL) receptor, apolipoproteins (e.g. apolipoprotein E or apolipoprotein A-I), IL-2 receptor and its antagonists, insulin, globin, immunoglobulins, catalytic antibodies, the interleukins, insulin-like growth factors, superoxide dismutase, immune responder modifiers, parathyroid hormone, interferons, nerve growth factors, tissue plasminogen activators, and colony stimulating factors, and variants of these proteins which have improved or novel biological properties or more desirable half-life or turnover times in vivo. Such a construct is also useful to integrate exogenous DNA (at a preselected site in a recipient cell's genome) which is a therapeutic product, such as DNA sequences sufficient for sequestration of a protein or nucleic acid in the transfected primary or secondary cell,

-21-

DNA sequences which bind to a cellular regulatory protein, DNA sequences which alter the secondary or tertiary chromosomal structure and DNA sequences which are transcriptional regulatory elements into genomic DNA of primary or
5 secondary cells.

The exogenous DNA, targeting sequences and selectable marker can be introduced into cells on a single DNA construct or on separate constructs. The total length of the DNA construct will vary according to the number of components (exogenous DNA, targeting sequences, selectable
10 marker gene) and the length of each. The entire construct length will generally be at least 20 nucleotides. In a construct in which the exogenous DNA has sufficient homology with genomic DNA to undergo homologous recombination,
15 the construct will include a single component, the exogenous DNA. In this embodiment, the exogenous DNA, because of its homology, serves also to target integration into genomic DNA and additional targeting sequences are unnecessary. Such a construct is useful to knock out, replace
20 or repair a resident DNA sequence, such as an entire gene, a gene portion, a regulatory element or portion thereof or regions of DNA which, when removed, place regulatory and structural sequences in functional proximity. It is also useful when the exogenous DNA contains a marker useful for
25 selection or amplification of linked sequences.

In a second embodiment, the DNA construct includes exogenous DNA, targeting DNA sequences and DNA encoding at least one selectable marker. In this second embodiment, the order of construct components can be: targeting sequences-exogenous DNA-DNA encoding a selectable marker(s)-targeting sequences. In this embodiment, one or
30 more selectable markers are included in the construct, which makes selection based on a selectable phenotype possible. Cells that stably integrate the construct will

-22-

survive treatment with the selective agent; a subset of the stably transfected cells will be HR cells, which can be identified by a variety of techniques, including PCR, Southern hybridization and phenotypic screening.

5 In a third embodiment, the order of components in the DNA construct can be: targeting sequence-selectable marker 1 - targeting sequence - selectable marker 2. In this embodiment selectable marker 2 displays the property of negative selection. That is, the gene product of selectable marker 2 can be selected against by growth in an appropriate media formulation containing an agent (typically a drug or metabolite analog) which kills cells expressing selectable marker 2. Recombination between the targeting sequences flanking selectable marker 1 with homologous sequences in the host cell genome results in the targeted integration of selectable marker 1, while selectable marker 2 is not integrated. Such recombination events generate cells which are stably transfected with selectable marker 1 but not stably transfected with selectable marker 2, and such cells can be selected for by growth in the media containing the selective agent which selects for selectable marker 1 and the selective agent which selects against selectable marker 2.

15 In a fourth embodiment, the DNA construct includes a positively selectable marker that allows for the selection of cells containing amplified copies of that marker. The amplification of such a marker results in the co-amplification of flanking DNA sequences. In this embodiment, the order of construct components is, for example: targeting sequences - DNA encoding an amplifiable positively selectable marker - DNA encoding a second selectable marker (optional) - DNA sequences corresponding to either an exogenous gene to be expressed under the control of a

-23-

suitable promoter or a promoter only which is positioned to activate an endogenous gene - targeting DNA sequences.

A DNA construct can include an inducible promoter which controls expression of the exogenous DNA, making
5 inducible expression possible. Optionally, the DNA construct can include a bacterial origin of replication and bacterial antibiotic resistance markers, which allow for large-scale plasmid propagation in bacteria. A DNA construct which includes DNA encoding a selectable marker,
10 along with additional sequences, such as a promoter, polyadenylation site and splice junctions, can be used to confer a selectable phenotype upon transfected primary or secondary cells (e.g., plasmid pCDNEO, schematically represented in Figure 2). Such a DNA construct can be
15 co-transfected into primary or secondary cells, along with a targeting DNA sequence, using methods described herein.

In all embodiments of the DNA construct, exogenous DNA can encode one or more products, can be one or more therapeutic products or one or more of each, thus making
20 it possible to deliver multiple products.

Uses of Transfected Cells

Cells produced using the methods and DNA constructs described herein can be used for a wide variety of purposes. Primary, secondary, or immortalized cells of vertebrate origin can be produced in which 1) DNA already
25 present in a recipient cell is repaired, altered, deleted, or replaced; 2) a gene or DNA sequence which encodes a therapeutic product (or other desired product) or is itself a therapeutic product is introduced into the genome
30 of a recipient cell at a preselected site (i.e. gene targeting); 3) regulatory sequences present in the primary, secondary or immortalized cell recipients have been repaired, altered, deleted or replaced; or 4) an entire gene

-24-

or gene portion has been repaired, altered, deleted, or replaced. Homologous recombination can also be used to produce universal donor cells, in which cell surface markers involved in histocompatibility have been altered, 5 deleted or replaced, or in which the expression of such markers is altered, impaired, or eliminated.

The cells of the present invention are useful for in vitro production of therapeutic products which can be purified and delivered by conventional pharmaceutical 10 routes. For example, primary, secondary, or immortalized human cells can be transfected with exogenous DNA containing a regulatory region which, upon homologous recombination with genomic DNA sequences, results in the replacement of an endogenous target gene's regulatory region with 15 a regulatory region that allows novel expression and/or regulation of the target gene and, ultimately, production of a therapeutically useful product by the transfected cell. The activated endogenous target gene can further be amplified if an appropriate selectable marker gene is 20 included in the targeting DNA. With or without amplification, these cells can be subject to large-scale cultivation for harvest of intracellular or extracellular protein products.

Transfected cells of the present invention are useful, as populations of transfected primary cells, transfected clonal cell strains, transfected heterogeneous cell strains, and as cell mixtures in which at least one representative cell of one of the three preceding categories of transfected cells is present, as a delivery system for 25 treating an individual with an abnormal or undesirable condition which responds to delivery of a therapeutic product, which is either: 1) a therapeutic protein (e.g., a protein which is absent, underproduced relative to the individual's physiologic needs, defective or inefficiently 30

-25-

or inappropriately utilized in the individual; a protein with novel functions, such as enzymatic or transport functions) or 2) a therapeutic nucleic acid (e.g., DNA which binds to or sequesters a regulatory protein, RNA which inhibits gene expression or has intrinsic enzymatic activity). In the method of the present invention of providing a therapeutic protein or nucleic acid, transfected primary cells, clonal cell strains or heterogenous cell strains are administered to an individual in whom the abnormal or undesirable condition is to be treated or prevented, in sufficient quantity and by an appropriate route, to express or make available the exogenous DNA at physiologically relevant levels. A physiologically relevant level is one which either approximates the level at which the product is produced in the body or results in improvement of the abnormal or undesirable condition. Cells administered in the present method are cells transfected with exogenous DNA which encodes a therapeutic product, exogenous DNA which is itself a therapeutic product or exogenous DNA, such as a regulatory sequence, which is introduced into a preselected site in genomic DNA through homologous recombination and functions to cause recipient cells to produce a product which is normally not expressed in the cells or to produce the product of a higher level than occurs in the corresponding nontransfected cell. In the embodiment in which a regulatory sequence (e.g., a promoter) is introduced, it replaces or disables a regulatory sequence normally associated with a gene, and results in expression of the gene at a higher level than occurs in the corresponding nontransfected cell or allows a pattern of regulation or induction that is different from the corresponding nontransfected cell.

Immortalized cells which produce a therapeutic protein produced by the methods described herein and

-26-

can be used in gene therapy whether made by cells produced by: 1) random integration of the therapeutic protein, 2) homologous recombination to target the therapeutic protein into a cell's genome, 3) homologous recombination to activate or turn on a gene of therapeutic interest, or 4) gene amplification in conjunction with one of the three preceding methods. According to the invention described herein, the immortalized cells are enclosed in one of a number of semipermeable barrier devices. The permeability properties of the device are such that the cells are prevented from leaving the device upon implantation into an animal, but the therapeutic product is freely permeable and can leave the barrier device and enter the local space surrounding the implant or enter the systemic circulation. A number of filtration membranes can be used for this purpose, including, but not limited to, cellulose, cellulose acetate, nitrocellulose, polysulfone, polyvinylidene difluoride, polyvinyl chloride polymers and polymers of polyvinyl chloride derivatives. Alternatively, barrier devices can be utilized to allow primary, secondary, or immortalized cells from another species to be used for gene therapy in humans. The use of cells from other species can be desirable in cases where the non-human cells are advantageous for protein production purposes or in cases where the non-human protein is therapeutically useful, for example, the use of cells derived from salmon for the production of salmon calcitonin and the use of cells derived from pigs for the production of porcine insulin.

Cells from non-human species can also be used for in vitro protein production. These cells can be immortalized, primary, or secondary cells which produce a therapeutic protein produced by the methods described here and in the U.S. patent applications incorporated herein by

-27-

peutic protein produced by the methods described here and in the U.S. patent applications incorporated her in by reference, whether made by cells produced by: 1) random integration of the therapeutic protein, 2) homologous recombination to target the therapeutic protein into a cell's genome, 3) homologous recombination to activate or turn on a gene of therapeutic interest, or 4) gene amplification in conjunction with one of the three preceding methods. The use of cells from other species may be desirable in cases where the non-human cells are advantageous for protein production purposes (for example CHO cells) or in cases where the non-human protein is therapeutically or commercially useful, for example, the use of cells derived from salmon for the production of salmon calcitonin, the use of cells derived from pigs for the production of porcine insulin, and the use of bovine cells for the production of bovine growth hormone.

Transfected cells of the present invention are useful in a number of applications in humans and animals. In one embodiment, the cells can be implanted into a human or an animal for protein delivery in the human or animal. For example, human growth hormone (hGH), human EPO (hEPO), or human insulinotropin can be delivered systemically in humans for therapeutic benefits. Barrier devices, through which the therapeutic product is freely permeable, can be used to retain cells in a fixed position in vivo or to protect and isolate the cells from the host's immune system. Barrier devices are particularly useful and allow transfected immortalized cells, transfected cells from another species (transfected xenogeneic cells), or cells from a nonhistocompatibility-matched donor (transfected allogeneic cells) to be implanted for treatment of human or animal conditions or for agricultural uses (i.e., meat and dairy production). Barrier devices also allow conve-

-28-

nient short-term (i.e., transient) therapy by providing ready access to the cells for removal when the treatment regimen is to be halted for any reason.

Transfected cells of the present invention are also
5 useful for eliciting antibody production or for immunizing humans and animals against pathogenic agents. Implanted transfected cells can be used to deliver immunizing anti-
gens that result in stimulation of the host's cellular and humoral immune responses. These immune responses can be
10 designed for protection of the host from future infectious agents (i.e., for vaccination), to stimulate and augment the disease-fighting capabilities directed against an ongoing infection, or to produce antibodies directed
against the antigen produced in vivo by the transfected
15 cells that can be useful for therapeutic or diagnostic purposes. Removable barrier devices can be used to allow a simple means of terminating exposure to the antigen. Alternatively, the use of cells that will ultimately be rejected (xenogeneic or allogeneic transfected cells) can
20 be used to limit exposure to the antigen since antigen production will cease when the cells have been rejected.

Explanation of the Examples

As described herein, Applicants have demonstrated that DNA can be introduced into primary, secondary or
25 immortalized vertebrate cells and integrated into the genome of the transfected primary or secondary cells by homologous recombination. That is, they have demonstrated gene targeting in primary, secondary and immortalized mammalian cells. They have further demonstrated that the
30 exogenous DNA has the desired function in the homologously recombinant (HR) cells and that correctly targeted cells can be identified on the basis of a detectable phenotype conferred by the properly targeted DNA.

-29-

In addition, the present invention relates to a method of protein production using transfected primary, secondary or immortalized cells. The method involves transfecting primary cells, secondary cells or immortal-
5 ized cells with exogenous DNA which encodes a therapeutic product or with DNA which is sufficient to target to and activate an endogenous gene which encodes a therapeutic product. For example, Examples 1g, 1j, 2, 3 and 4 describe protein production by targeting and activation of a
10 selected endogenous gene.

The applicants also describe DNA constructs and methods for amplifying an endogenous cellular gene that has been activated by gene targeting (Examples 1f-1k and Example 3) and further describe methods by which a gene
15 can be inserted at a preselected site in the genome of a primary, secondary, or immortalized cell by gene targeting (Example 1d).

Applicants describe construction of a plasmid useful for targeting to a particular locus (the HPRT locus) in
20 the human genome and selection based upon a drug resistant phenotype (Example 1a). This plasmid is designated pE3Neo and its integration into the cellular genomes at the HPRT locus produces cells which have an *hprt*⁺, 6-TG resistant phenotype and are also G418 resistant. As described, they
25 have shown that pE3Neo functions properly in gene targeting in an established human fibroblast cell line (Example 1b), by demonstrating localization of the DNA introduced into established cells within exon 3 of the HPRT gene.

In addition, Applicants demonstrate gene targeting in
30 primary and secondary human skin fibroblasts using pE3Neo (Example 1c). The subject application further demonstrates that modification of DNA termini enhances targeting of DNA into genomic DNA (Examples 1c and 1e).

-30-

Examples 1f-1h and 2 illustrate embodiments in which the normal regulatory sequences upstream of the human EPO gene are altered to allow expression of hEPO in primary or secondary fibroblast strains which do not express EPO in detectable quantities in their untransfected state. In one embodiment the product of targeting leaves the normal EPO protein intact, but under the control of the mouse metallothionein promoter. Examples 1i and 1j demonstrate the use of similar targeting constructs to activate the endogenous growth hormone gene in primary or secondary human fibroblasts. In other embodiments described for activating EPO expression in human fibroblasts, the products of targeting events are chimeric transcription units, in which the first exon of the human growth hormone gene is positioned upstream of EPO exons 2-5. The product of transcription (controlled by the mouse metallothionein promoter), splicing, and translation is a protein in which amino acids 1-4 of the hEPO signal peptide are replaced with amino acid residues 1-3 of hGH. The chimeric portion of this protein, the signal peptide, is removed prior to secretion from cells. Example 5 describes targeting constructs and methods for producing cells which will convert a gene (with introns) into an expressible cDNA copy of that gene (without introns) and the recovery of such expressible cDNA molecules in microbial (e.g., yeast or bacterial) cells. Example 6 describes construction of a targeting vector, designated PREPO4 for dual selection and selection of cells in which the dhfr gene is amplified. Plasmid PREPO4 has been used to amplify the human EPO (hEPO) locus in HT1080 cells (an immortalized human cell line) after activation of the endogenous hEPO gene by homologous recombination. As described, stepwise selection in methotrexate-containing media resulted in a 70-

-31-

fold increase in hEPO production in cells resistant to 0.4 μ M methotrexate.

The Examples provide methods for activating or for activating and amplifying endogenous genes by gene targeting which do not require manipulation or other uses of the target genes' protein coding regions. By these methods, normally inactive genes can be activated in cells that have properties desirable for *in vitro* protein production (e.g., pharmaceuticals) or *in vivo* protein delivery methods (e.g. gene therapy). Figures 5 and 6 illustrate two strategies for transcriptionally activating the hEPO gene.

Using the methods and DNA constructs or plasmids taught herein or modifications thereof which are apparent to one of ordinary skill in the art, exogenous DNA which encodes a therapeutic product (e.g., protein, ribozyme, nucleic acid) can be inserted at preselected sites in the genome of vertebrate (e.g., mammalian, both human and nonhuman) primary or secondary cells.

The present invention will now be illustrated by the following examples, which are not intended to be limiting in any way.

EXAMPLES

EXAMPLE 1. PRODUCTION OF TRANSFECTED CELL STRAINS BY GENE TARGETING

Gene targeting occurs when transfecting DNA either integrates into or partially replaces chromosomal DNA sequences through a homologous recombinant event. While such events can occur in the course of any given transfection experiment, they are usually masked by a vast excess of events in which plasmid DNA integrates by nonhomologous, or illegitimate, recombination.

-32-

a. GENERATION OF A CONSTRUCT USEFUL FOR SELECTION OF
GENE TARGETING EVENTS IN HUMAN CELLS

One approach to selecting the targeted events is by genetic selection for the loss of a gene function due to the integration of transfecting DNA. The human HPRT locus encodes the enzyme hypoxanthine-phosphoribosyl transferase. *hprt*⁻ cells can be selected for by growth in medium containing the nucleoside analog 6-thioguanine (6-TG): cells with the wild-type (HPRT⁺) allele are killed by 6-TG, while cells with mutant (*hprt*⁻) alleles can survive. Cells harboring targeted events which disrupt HPRT gene function are therefore selectable in 6-TG medium.

To construct a plasmid for targeting to the HPRT locus, the 6.9 kb HindIII fragment extending from positions 11,960-18,869 in the HPRT sequence (Genebank name HUMHPRTB; Edwards, A. *et al.*, *Genomics* 6:593-608 (1990)) and including exons 2 and 3 of the HPRT gene, is subcloned into the HindIII site of pUC12. The resulting clone is cleaved at the unique XhoI site in exon 3 of the HPRT gene fragment and the 1.1 kb SalI-XhoI fragment containing the neo gene from pMC1Neo (Stratagene) is inserted, disrupting the coding sequence of exon 3. One orientation, with the direction of neo transcription opposite that of HPRT transcription was chosen and designated pE3Neo. The replacement of the normal HPRT exon 3 with the neo-disrupted version will result in an *hprt*⁻, 6-TG resistant phenotype. Such cells will also be G418 resistant.

b. GENE TARGETING IN AN ESTABLISHED HUMAN FIBROBLAST
CELL LINE

As a demonstration of targeting in immortalized cell lines, and to establish that pE3Neo functions properly in gene targeting, the human fibrosarcoma cell line HT1080

-33-

(ATCC CCL 121) was transfected with pE3Neo by electroporation.

HT1080 cells were maintained in HAT (hypoxanthine/aminopterin/xanthine) supplemented DMEM with 15% calf serum (Hyclone) prior to electroporation. Two days before electroporation, the cells are switched to the same medium without aminopterin. Exponentially growing cells were trypsinized and diluted in DMEM/15% calf serum, centrifuged, and resuspended in PBS (phosphate buffered saline) at a final cell volume of 13.3 million cells per ml. pE3Neo is digested with HindIII, separating the 8 kb HPRT-neo fragment from the pUC12 backbone, purified by phenol extraction and ethanol precipitation, and resuspended at a concentration of 600 µg/ml. 50 µl (30 µg) was added to the electroporation cuvette (0.4 cm electrode gap; Bio-Rad Laboratories), along with 750 µl of the cell suspension (10 million cells). Electroporation was at 450 volts, 250 µFarads (Bio-Rad Gene Pulser; Bio-Rad Laboratories). The contents of the cuvette were immediately added to DMEM with 15% calf serum to yield a cell suspension of 1 million cells per 25 ml media. 25 ml of the treated cell suspension was plated onto 150 mm diameter tissue culture dishes and incubated at 37°C, 5% CO₂. 24 hrs later, a G418 solution was added directly to the plates to yield a final concentration of 800 µg/ml G418. Five days later the media was replaced with DMEM/15% calf serum/800 µg/ml G418. Nine days after electroporation, the media was replaced with DMEM/15% calf serum/800 µg/ml G418 and 10 µM 6-thioguanine. Colonies resistant to G418 and 6-TG were picked using cloning cylinders 14-16 days after the dual selection was initiated.

The results of five representative targeting experiments in HT1080 cells are shown in Table 1.

-34-

TABLE 1

Transfection	Number of Treated Cells	Number of G418 ^r 6-TG ^r Clones
1	1×10^7	32
2	1×10^7	28
3	1×10^7	24
4	1×10^7	32
5	1×10^7	66

For transfection 5, control plates designed to determine the overall yield of G418^r colonies indicated that 33,700 G418^r colonies could be generated from the initial 1×10^7 treated cells. Thus, the ratio of targeted to non-targeted events is 66/33,700, or 1 to 510. In the five experiments combined, targeted events arise at a frequency of 3.6×10^6 , or 0.00036% of treated cells.

Restriction enzyme and Southern hybridization experiments using probes derived from the neo and HPRT genes localized the neo gene to the HPRT locus at the predicted site within HPRT exon 3.

c. GENE TARGETING IN PRIMARY AND SECONDARY HUMAN SKIN FIBROBLASTS

pE3Neo is digested with HindIII, separating the 8 kb HPRT-neo fragment from the pUC12 backbone, and purified by phenol extraction and ethanol precipitation. DNA was resuspended at 2 mg/ml. Three million secondary human foreskin fibroblasts cells in a volume of 0.5 ml were electroporated at 250 volts and 960 μ Farads, with 100 μ g of HindIII pE3Neo (50 μ l). Three separate transfections were performed, for a total of 9 million treated cells.

-35-

Cells are processed and selected for G418 resistance. 500,000 cells per 150 mm culture dish were plated for G418 selection. After 10 days under selection, the culture medium is replaced with human fibroblast nutrient medium containing 400 µg/ml G418 and 10 µM 6-TG. Selection with the two drug combination is continued for 10 additional days. Plates are scanned microscopically to localize human fibroblast colonies resistant to both drugs. The fraction of G418^r t-TG^r colonies is 4 per 9 million treated cells. These colonies constitute 0.0001% (or 1 in a million) of all cells capable of forming colonies. Control plates designed to determine the overall yield of G418^r colonies indicated that 2,850 G418^r colonies could be generated from the initial 9×10^6 treated cells. Thus, the ratio of targeted to non-targeted events is 4/2,850, or 1 to 712. Restriction enzyme and Southern hybridization experiments using probes derived from the neo and HPRT genes were used to localize the neo gene to the HPRT locus at the predicted site within HPRT exon 3 and demonstrate that targeting had occurred in these four clonal cell strains. Colonies resistant to both drugs have also been isolated by transfecting primary cells ($1/3.0 \times 10^7$).

The results of several pE3Neo targeting experiments are summarized in Table 2. HindIII digested pE3Neo was either transfected directly or treated with exonuclease III to generate 5' single-stranded overhangs prior to transfection (see Example 1c). DNA preparations with single-stranded regions ranging from 175 to 930 base pairs in length were tested. Using pE3neo digested with HindIII alone, 1/799 G418-resistant colonies were identified by restriction enzyme and Southern hybridization analysis as having a targeted insertion of the neo gene at the HPRT locus (a total of 24 targeted clones were isolated).

-36-

Targeting was maximally stimulated (approximately 10-fold stimulation) when overhangs of 175 bp were used, with 1/80 G418^r colonies displaying restriction fragments that are diagnostic for targeting at HPRT (a total of 9 targeted clones were isolated). Thus, using the conditions and recombinant DNA constructs described here, targeting is readily observed in normal human fibroblasts and the overall targeting frequency (the number of targeted clones divided by the total number of clones stably transfected to G418-resistance) can be stimulated by transfection with targeting constructs containing single-stranded overhanging tails, by the method as described in Example 1e.

TABLE 2

TARGETING TO THE HPRT LOCUS IN HUMAN FIBROBLASTS

<u>pE3neo Treatment</u>	<u>Number of Experiments</u>	<u>Number Targeted Per G418^r Colony</u>	<u>Total Number of Targeted Clones</u>
HindIII digest	6	1/799	24
175 bp overhang	1	1/80	9
350 bp overhang	3	1/117	20
930 bp overhang	1	1/144	1

d. GENERATION OF A CONSTRUCT FOR TARGETED INSERTION OF A GENE OF THERAPEUTIC INTEREST INTO THE HUMAN GENOME AND ITS USE IN GENE TARGETING

A variant of pE3Neo, in which a gene of therapeutic interest is inserted within the HPRT coding region, adjacent to or near the neo gene, can be used to target a gene of therapeutic interest to a specific position in a recipient primary or secondary cell genome. Such a variant of

-37-

pE3Neo can be constructed for targeting the hGH gene to the HPRT locus.

pXGH5 (schematically presented in Figure 1) is digested with EcoRI and the 4.1 kb fragment containing the hGH gene and linked mouse metallothionein (mMT) promoter is isolated. The EcoRI overhangs are filled in with the Klenow fragment from *E. coli* DNA polymerase. Separately, pE3Neo is digested with XhoI, which cuts at the junction of the neo fragment and HPRT exon 3 (the 3' junction of the insertion into exon 3). The XhoI overhanging ends of the linearized plasmid are filled in with the Klenow fragment from *E. coli* DNA polymerase, and the resulting fragment is ligated to the 4.1 kb blunt-ended hGH-mMT fragment. Bacterial colonies derived from the ligation mixture are screened by restriction enzyme analysis for a single copy insertion of the hGH-mMT fragment and one orientation, the hGH gene transcribed in the same direction as the neo gene, is chosen and designated pE3Neo/hGH. pE3Neo/hGH is digested with HindIII, releasing the 12.1 kb fragment containing HPRT, neo and mMT-hGH sequences. Digested DNA is treated and transfected into primary or secondary human fibroblasts as described in Example 1c. G418^r TG^r colonies are selected and analyzed for targeted insertion of the mMT-hGH and neo sequences into the HPRT gene as described in Example 1c. Individual colonies are assayed for hGH expression using a commercially available immunoassay (Nichols Institute).

Secondary human fibroblasts were transfected with pE3Neo/hGH and thioguanine-resistant colonies were analyzed for stable hGH expression and by restriction enzyme and Southern hybridization analysis. Of thirteen TG^r colonies analyzed, eight colonies were identified with an insertion of the hGH gene into the endogenous HPRT locus. All eight strains stably expressed significant quantities

-38-

of hGH, with an average expression level of $22.7 \mu\text{g}/10^6$ cells/24 hours. Alternatively, plasmid pE3neoEPO, Figure 4, may be used to target EPO to the human HPRT locus.

The use of homologous recombination to target a gene of therapeutic interest to a specific position in a cell's genomic DNA can be expanded upon and made more useful for producing products for therapeutic purposes (e.g., pharmaceuticals, gene therapy) by the insertion of a gene through which cells containing amplified copies of the gene can be selected for by exposure of the cells to an appropriate drug selection regimen. For example, pE3neo/hGH (Example 1d) can be modified by inserting the dhfr, ada, or CAD gene at a position immediately adjacent to the hGH or neo genes in pE3neo/hGH. Primary, secondary, or immortalized cells are transfected with such a plasmid and correctly targeted events are identified. These cells are further treated with increasing concentrations of drugs appropriate for the selection of cells containing amplified genes (for dhfr, the selective agent is methotrexate, for CAD the selective agent is N-(phosphonacetyl)-L-aspartate (PALA), and for ada the selective agent is an adenine nucleoside (e.g., alanosine). In this manner the integration of the gene of therapeutic interest will be coamplified along with the gene for which amplified copies are selected. Thus, the genetic engineering of cells to produce genes for therapeutic uses can be readily controlled by preselecting the site at which the targeting construct integrates and at which the amplified copies reside in the amplified cells.

30 e. MODIFICATION OF DNA TERMINI TO ENHANCE TARGETING

Several lines of evidence suggest that 3'-overhanging ends are involved in certain homologous recombination pathways of E. coli, bacteriophage, S. cerevisiae and

-39-

Xenopus laevis. In *Xenopus laevis* oocytes, molecules with 3'-overhanging ends of several hundred base pairs in length underwent recombination with similarly treated molecules much more rapidly after microinjection than
5 molecules with very short overhangs (4 bp) generated by restriction enzyme digestion. In yeast, the generation of 3'-overhanging ends several hundred base pairs in length appears to be a rate limiting step in meiotic recombination. No evidence for an involvement of 3'-overhanging ends
10 in recombination in human cells has been reported, and in no case have modified DNA substrates of any sort been shown to promote targeting (one form of homologous recombination) in any species. In human cells, the effect of 3'-overhanging ends on targeting is untested. The experi-
15 ment described in the following example and Example 1c suggests that 5'-overhanging ends are effective for stimulating targeting in primary, secondary and immortalized human fibroblasts.

There have been no reports on the enhancement of
20 targeting by modifying the ends of the transfecting DNA molecules. This example serves to illustrate that modification of the ends of linear DNA molecules, by conversion of the molecules' termini from a double-stranded form to a single-stranded form, can stimulate targeting into the
25 genome of primary and secondary human fibroblasts.

1100 µg of plasmid pE3Neo (Example 1a) is digested with HindIII. This DNA can be used directly after phenol extraction and ethanol precipitation, or the 8 kb HindIII fragment containing only HPRT and the neo gene can be
30 separated away from the pUC12 vector sequences by gel electrophoresis. ExoIII digestion of the HindIII digested DNA results in extensive exonucleolytic digestion at each end, initiating at each free 3' end, and leaving 5'-overhanging ends. The extent of exonucleolytic action

-40-

and, hence, the length of the resulting 5'-overhangs, can be controlled by varying the time of ExoIII digestion. ExoIII digestion of 100 µg of HindIII digested pE3Neo is carried out according to the supplier's recommended conditions, for times of 30 sec, 1 min, 1.5 min, 2 min, 2.5 min, 3 min, 3.5 min, 4 min, 4.5 min, and 5 min. To monitor the extent of digestion an aliquot from each time point, containing 1 µg of ExoIII treated DNA, is treated with mung bean nuclease (Promega), under conditions recommended by the supplier, and the samples fractionated by gel electrophoresis. The difference in size between non-treated, HindIII digested pE3Neo and the same molecules treated with ExoIII and mung bean nuclease is measured. This size difference divided by two gives the average length of the 5'-overhang at each end of the molecule. Using the time points described above and digestion at 30°, the 5'-overhangs produced should range from 100 to 1,000 bases.

60 µg of ExoIII treated DNA (total HindIII digest of pE3Neo) from each time point is purified and electroporated into primary, secondary, or immortalized human fibroblasts under the conditions described in Example 1c. The degree to which targeting is enhanced by each ExoIII treated preparation is quantified by counting the number of G418^r 6-TG^r colonies and comparing these numbers to targeting with HindIII digested pE3Neo that was not treated with ExoIII.

The effect of 3'-overhanging ends can also be quantified using an analogous system. In this case HindIII digested pE3Neo is treated with bacteriophage T7 gene 6 exonuclease (United States Biochemicals) for varying time intervals under the supplier's recommended conditions. Determination of the extent of digestion (average length of 3'-overhang produced per end) and electroporation

-41-

conditions are as described for ExoIII treated DNA. The degree to which targeting is enhanced by each T7 gene 6 exonuclease treated preparation is quantified by counting the number of G418^r 6-TG^r colonies and comparing these numbers to targeting with HindIII digested pE3Neo that was not treated with T7 gene 6 exonuclease.

Other methods for generating 5' and 3' overhanging ends are possible, for example, denaturation and annealing of two linear molecules that partially overlap with each other will generate a mixture of molecules, each molecule having 3'-overhangs at both ends or 5'-overhangs at both ends, as well as reannealed fragments indistinguishable from the starting linear molecules. The length of the overhangs is determined by the length of DNA that is not in common between the two DNA fragments.

f. CONSTRUCTION OF TARGETING PLASMIDS FOR PLACING THE HUMAN ERYTHROPOIETIN GENE UNDER THE CONTROL OF THE MOUSE METALLOTHIONEIN PROMOTER IN PRIMARY, SECONDARY AND IMMORTALIZED HUMAN FIBROBLASTS

The following serves to illustrate one embodiment of the present invention, in which the normal positive and negative regulatory sequences upstream of the human erythropoietin (EPO) gene are altered to allow expression of human erythropoietin in primary, secondary or immortalized human fibroblasts, which do not express EPO in significant quantities as obtained.

A region lying exclusively upstream of the human EPO coding region can be amplified by PCR. Three sets of primers useful for this purpose were designed after analysis of the published human EPO [Genbank designation HUMERPA; Lin, F-K., et al., Proc. Natl. Acad. Sci., USA 82:7580-7584 (1985)]. These primer pairs can amplify fragments of 609, 603, or 590 bp.

-42-

TABLE 3

Primer	HUMERPA Coordinate	Sequence	Fragment Size
F1	2 - 20	5' AGCTTCTGGGCTTCCAGAC (SEQ ID NO 1)	
R2	610 - 595	5' GGGGTCCCTCAGCGAC (SEQ ID NO 2)	609 bp
F2	8 - 24	5' TGGGCTTCCAGACCCAG (SEQ ID NO 3)	
R2	610 - 595	5' GGGGTCCCTCAGCGAC	603 bp
F3	21 - 40	5' CCAGCTACTTTGCGGAATC (SEQ ID NO 4)	
R2	610 - 595	5' GGGGTCCCTCAGCGAC	590 bp

The three fragments overlap substantially and are interchangeable for the present purposes. The 609 bp fragment, extending from -623 to -14 relative to the translation start site (HUMERPA nucleotide positions 2 to 610), is ligated at both ends with ClaI linkers. The resulting ClaI-linked fragment is digested with ClaI and inserted into the ClaI site of pBluescriptIISK/+ (Stratagene), with the orientation such that HUMERPA nucleotide position 610 is adjacent to the SalI site in the plasmid polylinker). This plasmid, p5'EPO, can be cleaved, separately, at the unique FspI or SfiI sites in the EPO upstream fragment (HUMERPA nucleotide positions 150 and 405, respectively) and ligated to the mouse metallothionein promoter. Typically, the 1.8 kb EcoRI-BglII from the

-43-

mMT-I gene [containing no mMT coding sequences; Hamer, D.H. and Walling M., J. Mol. Appl. Gen. 1:273 288 (1982); this fragment can also be isolated by known methods from mouse genomic DNA using PCR primers designed from analysis of mMT sequences available from Genbank; i.e., MUSMTI, MUSMTIP, MUSMTIPRM] is made blunt-ended by known methods and ligated with SfiI digested (also made blunt-ended) or FspI digested p5'EPO. The orientations of resulting clones are analyzed and those in which the former mMT BglII site is proximal to the SalI site in the plasmid polylinker are used for targeting primary and secondary human fibroblasts. This orientation directs mMT transcription towards HUMERPA nucleotide position 610 in the final construct. The resulting plasmids are designated p5'EPO-mMTF and p5'EPO-mMTS for the mMT insertions in the FspI and SfiI sites, respectively.

Additional upstream sequences are useful in cases where it is desirable to modify, delete and/or replace negative regulatory elements or enhancers that lie upstream of the initial target sequence. In the case of EPO, a negative regulatory element that inhibits EPO expression in extrahepatic and extrarenal tissues [Semenza, G.L. et al., Mol. Cell. Biol. 10:930-938 (1990)] can be deleted. A series of deletions within the 6 kb fragment are prepared. The deleted regions can be replaced with an enhancer with broad host-cell activity [e.g. an enhancer from the Cytomegalovirus (CMV)].

The orientation of the 609 bp 5'EPO fragment in the pBluescriptIISK/+ vector was chosen since the HUMERPA sequences are preceded on their 5' end by a BamHI (distal) and HindIII site (proximal). Thus, a 6 kb BamHI-HindIII fragment normally lying upstream of the 609 bp fragment [Semenza, G. L. et al., Mol. Cell. Biol. 10:930-938 (1990)] can be isolated from genomic DNA by known methods.

-44-

For example, a bacteriophage, cosmid, or yeast artificial chromosome library could be screened with the 609 bp PCR amplified fragment as a probe. The desired clone will have a 6 kb BamHI-HindIII fragment and its identity can be confirmed by comparing its restriction map from a restriction map around the human EPO gene determined by known methods. Alternatively, constructing a restriction map of the human genome upstream of the EPO gene using the 609 bp fragment as a probe can identify enzymes which generate a fragment originating between HUMERPA coordinates 2 and 609 and extending past the upstream BamHI site; this fragment can be isolated by gel electrophoresis from the appropriate digest of human genomic DNA and ligated into a bacterial or yeast cloning vector. The correct clone will hybridize to the 609 bp 5'EPO probe and contain a 6 kb BamHI-HindIII fragment. The isolated 6 kb fragment is inserted in the proper orientation into p5'EPO, p5'EPO-mMTF, or p5'EPO-mMTS (such that the HindIII site is adjacent to HUMERPA nucleotide position 2). Additional upstream sequences can be isolated by known methods, using chromosome walking techniques or by isolation of yeast artificial chromosomes hybridizing to the 609 bp 5'EPO probe.

The cloning strategies described above allow sequences upstream of EPO to be modified in vitro for subsequent targeted transfection of primary, secondary or immortalized human fibroblasts. The strategies describe simple insertions of the mMT promoter, as well as deletion of the negative regulatory region, and deletion of the negative regulatory region and replacement with an enhancer with broad host-cell activity.

-45-

g. TARGETING TO SEQUENCES FLANKING THE HUMAN EPO GENE
AND ISOLATION OF TARGETED PRIMARY, SECONDARY AND
IMMORTALIZED HUMAN FIBROBLASTS BY SCREENING

For targeting, the plasmids are cut with restriction enzymes that free the insert away from the plasmid backbone. In the case of p5'EPO-mMTS, HindIII and SalI digestion releases a targeting fragment of 2.4 kb, comprised of the 1.8 kb mMT promoter flanked on the 5' and 3' sides by 405 bp and 204 base pairs, respectively, of DNA for targeting this construct to the regulatory region of the EPO gene. This DNA or the 2.4 kb targeting fragment alone is purified by phenol extraction and ethanol precipitation and transfected into primary or secondary human fibroblasts under the conditions described in Example 1c.

Transfected cells are plated onto 150 mm dishes in human fibroblast nutrient medium. 48 hours later the cells are plated into 24 well dishes at a density of 10,000 cells/cm² [approximately 20,000 cells per well; if targeting occurs at a rate of 1 event per 10⁶ clonable cells (Example 1c, then about 50 wells would need to be assayed to isolate a single expressing colony]. Cells in which the transfecting DNA has targeted to the homologous region upstream of EPO will express EPO under the control of the mMT promoter. After 10 days, whole well supernatants are assayed for EPO expression using a commercially available immunoassay kit (Amgen). Clones from wells displaying EPO synthesis are isolated using known methods, typically by assaying fractions of the heterogenous populations of cells separated into individual wells or plates, assaying fractions of these positive wells, and repeating as needed, ultimately isolating the targeted colony by screening 96-well microtiter plates seeded at one cell per well. DNA from entire plate lysates can also be analyzed by PCR for amplification of a fragment using a mMT specific

-46-

primer in conjunction with a primer lying upstream of HUMERPA nucleotide position 1. This primer pair should amplify a DNA fragment of a size precisely predicted based on the DNA sequence. Positive plates are trypsinized and replated at successively lower dilutions, and the DNA preparation and PCR steps repeated as needed to isolate targeted cells.

The targeting schemes herein described can also be used to activate hGH expression in immortalized human cells (for example, HT1080 (ATCC CCL 121) fibroblasts, HeLa (ATCC CCL 2) cells, MCF-7 (ATCC HTB 22) breast cancer cells, K-562 (ATCC CCL 243) leukemia cells, KB (ATCC CCL 17) carcinoma cells or 2780AD (Van der Bliek, A.M. *et al.*, Cancer Res. 48:5927-5932 (1988)) ovarian carcinoma cells) for the purposes of producing hGH for conventional pharmaceutical delivery.

h. TARGETING TO SEQUENCES FLANKING THE HUMAN EPO GENE AND ISOLATION OF TARGETED PRIMARY, SECONDARY AND IMMORTALIZED HUMAN FIBROBLASTS BY A POSITIVE OR A COMBINED POSITIVE/NEGATIVE SELECTION SYSTEM

The strategy for constructing p5'EPO-mMTF, p5'EPO-mMTS, and derivatives of such with the additional upstream 6 kb BamHI-HindIII fragment can be followed with the additional step of inserting the neo gene adjacent to the mMT promoter. In addition, a negative selection marker, for example, gpt [from pMSG (Pharmacia) or another suitable source], can be inserted adjacent to the HUMERPA sequences in the pBluescriptIISK/+ polylinker. In the former case, G418^r colonies are isolated and screened by PCR amplification or restriction enzyme and Southern hybridization analysis of DNA prepared from pools of colonies to identify targeted colonies. In the latter case, G418^r colonies are placed in medium containing 6-thioxanthine to select

-47-

against the integration of the gpt gene [Besnard, C. et al., Mol. Cell. Biol. 7:4139-4141 (1987)]. In addition, the HSV-TK gene can be placed on the opposite side of the insert as gpt, allowing selection for neo and against both gpt and TK by growing cells in human fibroblast nutrient medium containing 400 µg/ml G418, 100 µM 6-thioxanthine, and 25 µg/ml gancyclovir. The double negative selection should provide a nearly absolute selection for true targeted events and Southern blot analysis provides an ultimate confirmation.

The targeting schemes herein described can also be used to activate hEPO expression in immortalized human cells (for example, HT1080 fibroblasts, HeLa cells, MCF-7 breast cancer cells, K-562 leukemia cells, KB carcinoma cells or 2780AD ovarian carcinoma cells) for the purposes of producing hEPO for conventional pharmaceutical delivery.

i. CONSTRUCTION OF TARGETING PLASMIDS FOR PLACING THE HUMAN GROWTH HORMONE GENE UNDER THE CONTROL OF THE MOUSE METALLOTHIONEIN PROMOTER IN PRIMARY, SECONDARY OR IMMORTALIZED HUMAN FIBROBLASTS

The following example serves to illustrate one embodiment of the present invention, in which the normal regulatory sequences upstream of the human growth hormone gene are altered to allow expression of human growth hormone in primary, secondary or immortalized human fibroblasts.

Targeting molecules similar to those described in Example 1f for targeting to the EPO gene regulatory region are generated using cloned DNA fragments derived from the 5' end of the human growth hormone N gene. An approximately 1.8 kb fragment spanning HUMGHCSA (Genbank Entry nucleotide positions 3787-5432 (the positions of two EcoNI sites which generate a convenient sized fragment for

-48-

cloning or for diagnostic digestion of subclones involving this fragment) is amplified by PCR primers designed by analysis of the HUMGHCSA sequence in this region. This region extends from the middle of hGH gene N intron 1 to an upstream position approximately 1.4 kb 5' to the translational start site. pUC12 is digested with EcoRI and BamHI, treated with Klenow to generate blunt ends, and recircularized under dilute conditions, resulting in plasmids which have lost the EcoRI and BamHI sites. This plasmid is designated pUC12XEB. HindIII linkers are ligated onto the amplified hGH fragment and the resulting fragment is digested with HindIII and ligated to HindIII digested pUC12XEB. The resulting plasmid, pUC12XEB-5'hGH, is digested with EcoRI and BamHI, to remove a 0.5 kb fragment lying immediately upstream of the hGH transcriptional initiation site. The digested DNA is ligated to the 1.8 kb EcoRI-BglII from the mMT-I gene [containing no mMT coding sequences; Hamer, D.H. and Walling, M., J. Mol. Appl. Gen. 1:273-288 (1982); the fragment can also be isolated by known methods from mouse genomic DNA using PCR primers designed from analysis of mMT sequences available from Genbank; i.e., MUSMTI, MUSMTIP, MUSMTIPRM]. This plasmid p5'hGH-mMT has the mMT promoter flanked on both sides by upstream hGH sequences.

The cloning strategies described above allow sequences upstream of hGH to be modified in vitro for subsequent targeted transfection of primary, secondary or immortalized human fibroblasts. The strategy described a simple insertion of the mMT promoter. Other strategies can be envisioned, for example, in which an enhancer with broad host-cell specificity is inserted upstream of the inserted mMT sequence.

-49-

j. TARGETING TO SEQUENCES FLANKING THE HUMAN hGH GENE
AND ISOLATION OF TARGETED PRIMARY, SECONDARY AND
IMMORTALIZED HUMAN FIBROBLASTS BY SCREENING

For targeting, the plasmids are cut with restriction enzymes that free the insert away from the plasmid backbone. In the case of p5'hGH-mMT, HindIII digestion releases a targeting fragment of 2.9 kb, comprised of the 1.8 kb mMT promoter flanked on the 5' end 3' sides by DNA for targeting this construct to the regulatory region of the hGH gene. This DNA or the 2.9 kb targeting fragment alone is purified by phenol extraction and ethanol precipitation and transfected into primary or secondary human fibroblasts under the conditions described in Example 11. Transfected cells are plated onto 150 mm dishes in human fibroblast nutrient medium. 48 hours later the cells are plated into 24 well dishes at a density of 10,000 cells/cm² [approximately 20,000 cells per well; if targeting occurs at a rate of 1 event per 10⁶ clonable cells (Example 1c), then about 50 wells would need to be assayed to isolate a single expressing colony]. Cells in which the transfecting DNA has targeted to the homologous region upstream of hGH will express hGH under the control of the mMT promoter. After 10 days, whole well supernatants are assayed for hGH expression using a commercially available immunoassay kit (Nichols). Clones from wells displaying hGH synthesis are isolated using known methods, typically by assaying fractions of the heterogenous populations of cells separated into individual wells or plates, assaying fractions of these positive wells, and repeating as needed, ultimately isolated the targeted colony by screening 96-well microtiter plates seeded at one cell per well. DNA from entire plate lysates can also be analyzed by PCR for amplification of a fragment using a mMT specific primer in conjunction with a primer lying downstream of

-50-

HUMGHCSA nucleotide position 5,432. This primer pair should amplify a DNA fragment of a size precisely predicted based on the DNA sequence. Positive plates are trypsinized and replated at successively lower dilutions, and
5 the DNA preparation and PCR steps repeated as needed to isolate targeted cells.

The targeting schemes herein described can also be used to activate hGH expression in immortalized human cells (for example, HT1080 fibroblasts, HeLa cells, MCF-7
10 breast cancer cells, K-562 leukemia cells, KB carcinoma cells or 2780AD ovarian carcinoma cells) for the purposes of producing hGH for conventional pharmaceutical delivery.

k. TARGETING TO SEQUENCES FLANKING THE HUMAN hGH GENE
AND ISOLATION OF TARGETED PRIMARY, SECONDARY AND
15 IMMORTALIZED HUMAN FIBROBLASTS BY A POSITIVE OR A
COMBINED POSITIVE/NEGATIVE SELECTION SYSTEM

The strategy for constructing p5'hGH-mMT can be followed with the additional step of inserting the neo gene adjacent to the mMT promoter. In addition, a negative selection marker, for example, gpt [from pMSG (Phar-
20 macia) or another suitable source], can be inserted adjacent to the HUMGHCSA sequences in the pUC12 poly-linker. In the former case, G418^r colonies are isolated and screened by PCR amplification or restriction enzyme and
25 Southern hybridization analysis of DNA prepared from pools of colonies to identify targeted colonies. In the latter case, G418^r colonies are placed in medium containing thioxanthine to select against the integration of the gpt gene (Besnard, C. et al., Mol. Cell. Biol. 7: 4139-4141
30 (1987)]. In addition, the HSV-TK gene can be placed on the opposite side of the insert as gpt, allowing selection for neo and against both gpt and TK by growing cells in human fibroblast nutrient medium containing 400 µg/ml

-51-

G418, 100 μ M 6-thioxanthine, and 25 μ g/ml gancyclovir. The double negative selection should provide a nearly absolute selection for true targeted events. Southern hybridization analysis is confirmatory.

5 The targeting schemes herein described can also be used to activate hGH expression in immortalized human cells (for example, HT1080 fibroblasts, HeLa cells, MCF-7 breast cancer cells, K-562 leukemia cells, KB carcinoma cells or 2780AD ovarian carcinoma cells) for the purposes
10 of producing hGH for conventional pharmaceutical delivery.

 The targeting constructs described in Examples 1f and 1i, and used in Examples 1g, 1h, 1j and 1k can be modified to include an amplifiable selectable marker (e.g., ada, dhfr, or CAD) which is useful for selecting cells in which
15 the activated endogenous gene, and the amplifiable selectable marker, are amplified. Such cells, expressing or capable of expressing the endogenous gene encoding a therapeutic product can be used to produce proteins (e.g., hGH and hEPO) for conventional pharmaceutical delivery or
20 for gene therapy.

1. TRANSFECTION OF PRIMARY AND SECONDARY FIBROBLASTS WITH EXOGENOUS DNA AND A SELECTABLE MARKER GENE BY ELECTROPORATION

 Exponentially growing or early stationary phase
25 fibroblasts are trypsinized and rinsed from the plastic surface with nutrient medium. An aliquot of the cell suspension is removed for counting, and the remaining cells are subjected to centrifugation. The supernatant is aspirated and the pellet is resuspended in 5 ml of elec-
30 troporation buffer (20 mM HEPES pH 7.3, 137 mM NaCl, 5 mM KCl, 0.7 mM Na_2HPO_4 , 6 mM dextrose). The cells are recentrifuged, the supernatant aspirated, and the cells resuspended in electroporation buffer containing 1 mg/ml acetylated bovine serum albumin. The final cell suspension

-52-

contains approximately 3×10^6 cells/ml. Electroporation should be performed immediately following resuspension.

- Supercoiled plasmid DNA is added to a sterile cuvette with a 0.4 cm electrode gap (Bio-Rad.) The final DNA concentration is generally at least 120 $\mu\text{g/ml}$. 0.5 ml of the cell suspension (containing approximately 1.5×10^6 cells) is then added to the cuvette, and the cell suspension and DNA solutions are gently mixed. Electroporation is performed with a Gene-Pulser apparatus (Bio-Rad). Capacitance and voltage are set at 960 μF and 250-300 V, respectively. As voltage increases, cell survival decreases, but the percentage of surviving cells that stably incorporate the introduced DNA into their genome increases dramatically. Given these parameters, a pulse time of approximately 14-20 mSec should be observed.

- Electroporated cells are maintained at room temperature for approximately 5 min, and the contents of the cuvette are then gently removed with a sterile transfer pipette. The cells are added directly to 10 ml of prewarmed nutrient media (as above with 15% calf serum) in a 10 cm dish and incubated as described above. The following day, the media is aspirated and replaced with 10 ml of fresh media and incubated for a further 16-24 hours. Subculture of cells to determine cloning efficiency and to select for G418-resistant colonies is performed the following day. Cells are trypsinized, counted and plated; typically, fibroblasts are plated at 10^3 cells/10 cm dish for the determination of cloning efficiency and at $1-2 \times 10^4$ cells/10 cm dish for G418 selection.

- Human fibroblasts are selected for G418 resistance in medium consisting of 300-400 $\mu\text{g/ml}$ G418 (Geneticin, disulfate salt with a potency of approximately 50%; Gibco) in fibroblasts nutrient media (with 15% calf serum). Cloning efficiency is determined in the absence of G418. The

-53-

plated cells are incubated for 12-14 days, at which time colonies are fixed with formalin, stained with crystal violet and counted (for cloning efficiency plated) or isolated using cloning cylinders (for G418 plates).

- 5 Electroporation and selection of rabbit fibroblasts is performed essentially as described for human fibroblasts, with the exception of the selection conditions used. Rabbit fibroblasts are selected for G418 resistance in medium containing 1 gm/ml G418.

- 10 Fibroblasts were isolated from freshly excised human foreskins. Cultures were seeded at 50,000 cells/cm in DMEM + 10% calf serum. When cultures became confluent, fibroblasts were harvested by trypsinization and transfected by electroporation. Electroporation conditions
15 were evaluated by transfection with the plasmid pCDNEO. A representative electroporation experiment using near optimal conditions (60 µg of plasmid pCDNEO at an electroporation voltage of 250 volts and a capacitance setting of 960 µFarads) resulted in one G418 colony per 588 treated
20 cells (0.17% of all cells treated), or one G418 colony per 71 clonable cells (1.4%).

- When nine separate electroporation experiments at near optimal conditions (60 µg of plasmid pCDNEO at an electroporation voltage of 300 volts and a capacitance
25 setting of 960 µFarads) were performed, an average of one G418 colony per 1,899 treated cells (0.05%) was observed, with a range of 1/882 to 1/7,500 treated cells. This corresponds to an average of one G418 colony per 38 clonable cells (2.6%).

- 30 Low passage primary human fibroblasts were converted to hGH expressing cells by co-transfection with plasmids; pCDNEO and pXGH5. Typically, 60 µg of an equimolar mixture of the two plasmids were transfected at near optimal conditions (electroporation voltage of 300 volts and a

-54-

capacitance setting of 960 μ Farads). The results of such an experiment resulted in one G418 colony per 14,705 treated cells.

hGH expression data for these and other cells isolated under identical transfection conditions are summarized below. Ultimately, 98% of all G418^r colonies could be expanded to generate mass cultures.

	Number of G418 ^r Clones	
	Analyzed	154
10	Number of G418 ^r /hGH	
	Expressing Clones	65
	Average hGH Expression	
	Level	2.3 μ g hGH/ 10^6 Cells/24 hr
	Maximum hGH Expression	
15	Level	23.0 μ g hGH/ 10^6 Cells/24 hr

Stable transfectants also have been generated by electroporation of primary or secondary human fibroblasts with pXGH301, a DNA construct in which the neo and hGH genes are present on the same plasmid molecule. pXGH301 was constructed by a two-step procedure. The SaII-ClaI fragment from pBR322 (positions 23-651 in pBR322) was isolated and inserted into SaII-ClaI digested pCDNEO, introducing a BamHI site upstream of the SV40 early promoter region of pCDNEO. This plasmid, pBNEO was digested with BamHI and the 2.1 kb fragment containing the neo gene under the control of the SV40 early promoter, was isolated and inserted into BamHI digested pXGH5. A plasmid with a single insertion of the 2.1 kb BamHI fragment was isolated in which neo and hGH are transcribed in the same direction relative to each other. This plasmid was designated pXGH301. For example, 1.5×10^6 cells were electroporated with 60 μ g pXGH301 at 300 volts and 960 μ Farads. G418

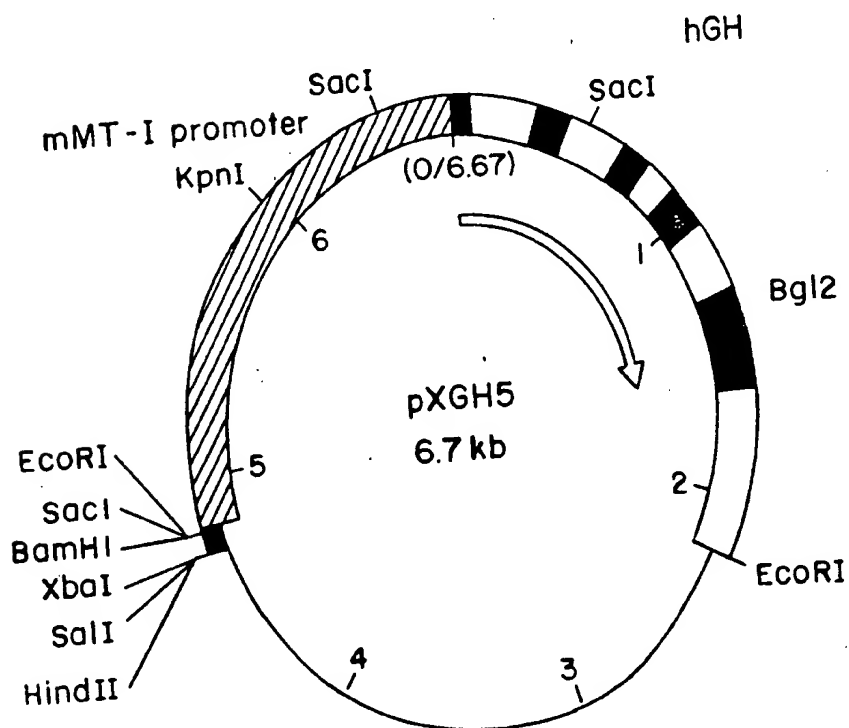


FIG. 1

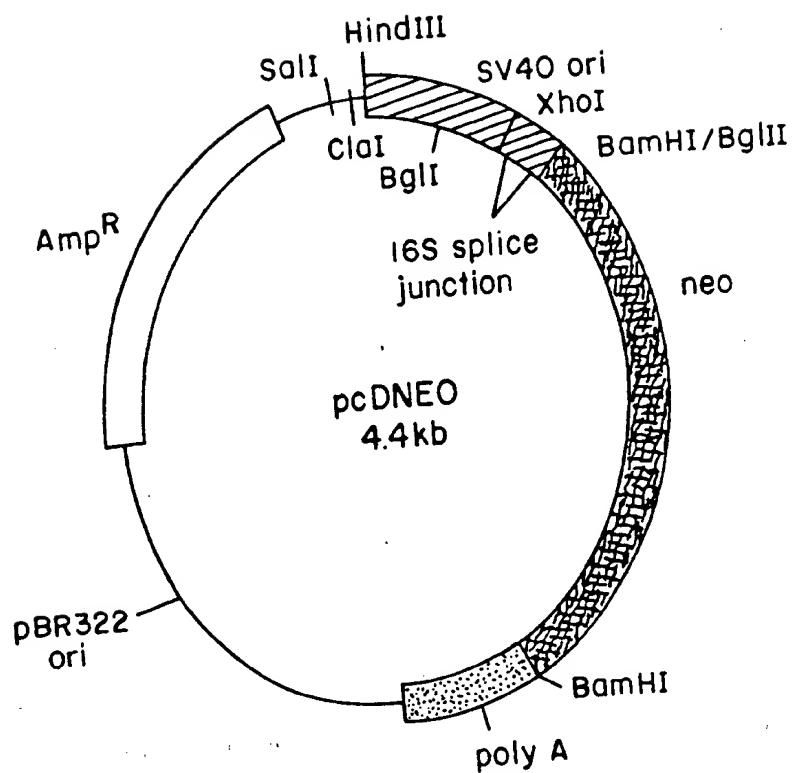


FIG. 2

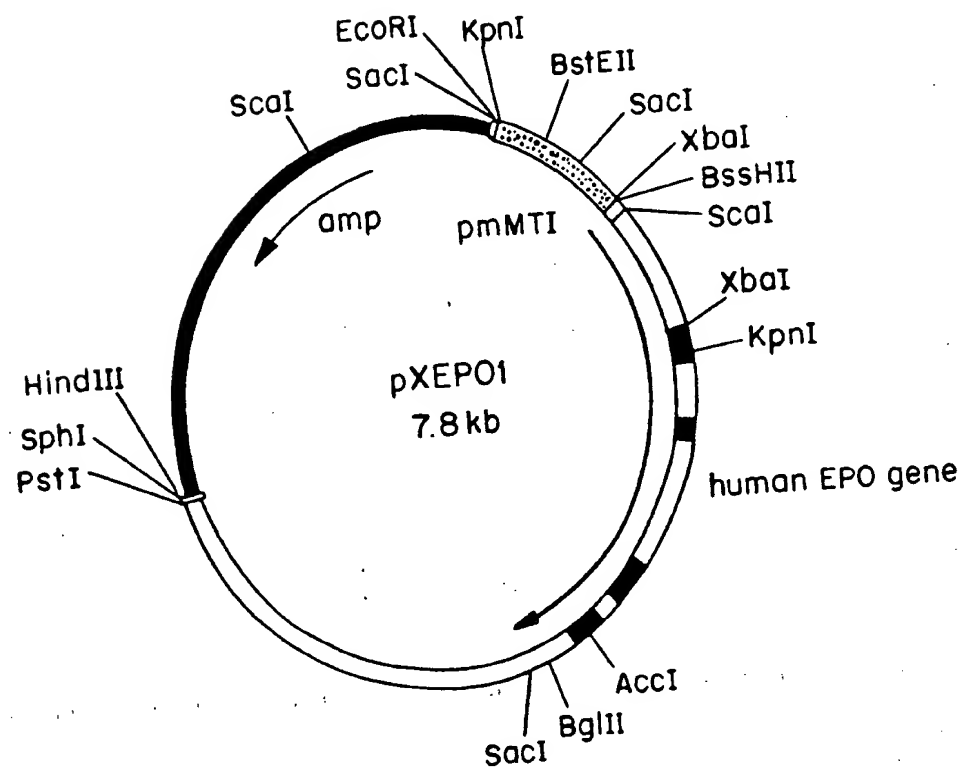


FIG. 3

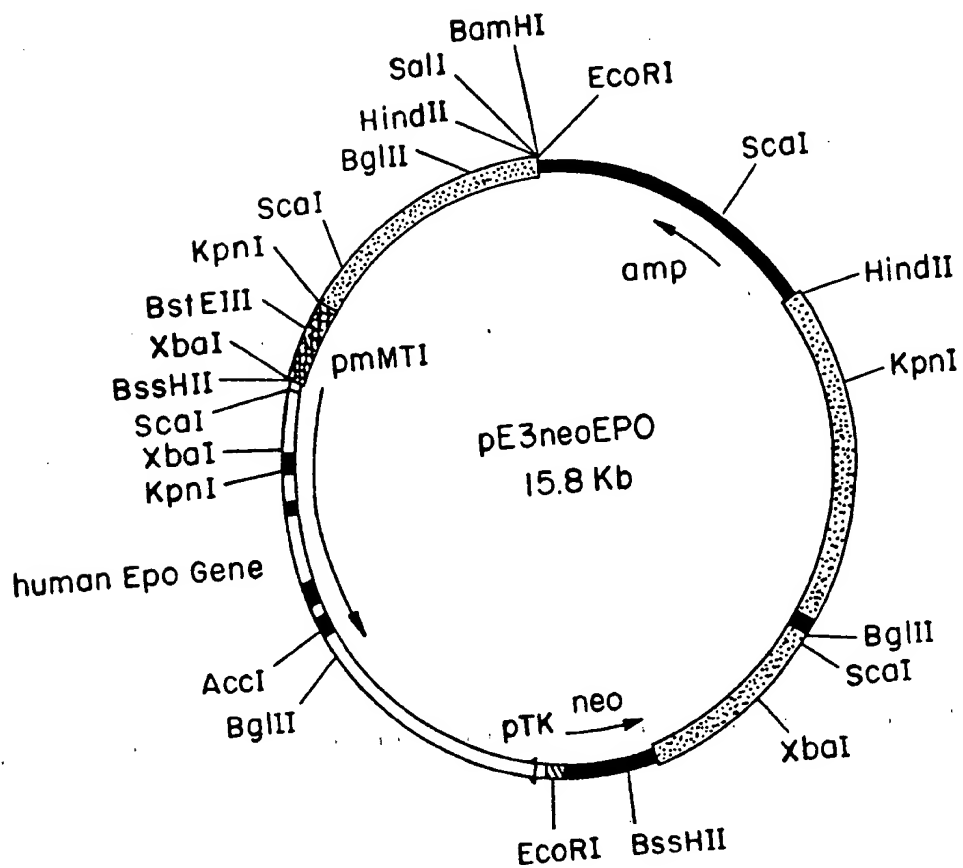


FIG. 4

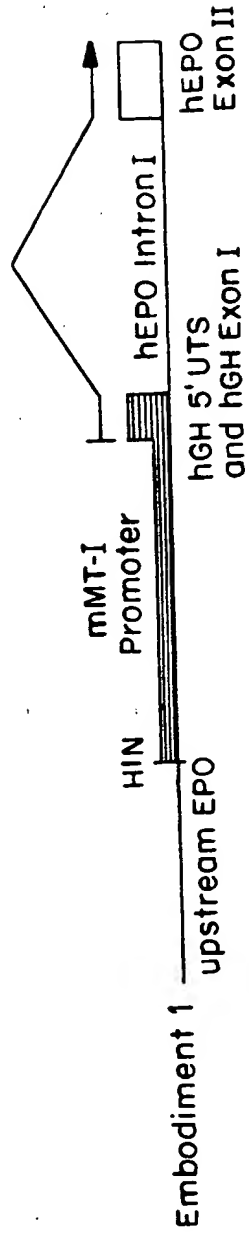


FIG. 5

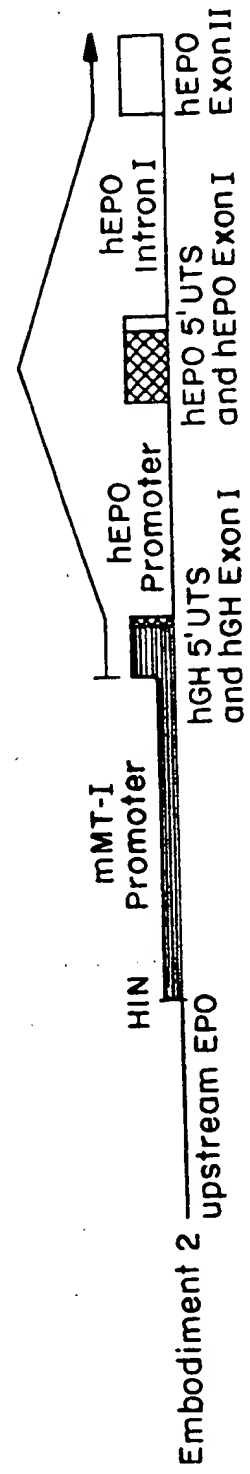


FIG. 6

6/7

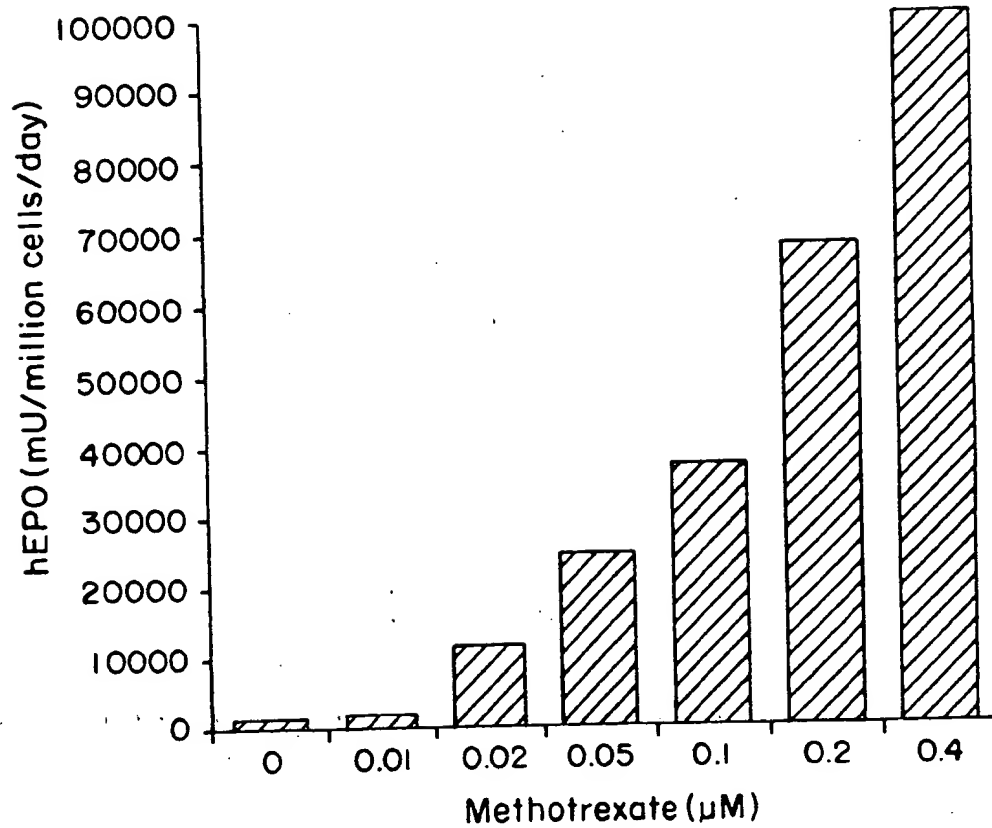


FIG. 7

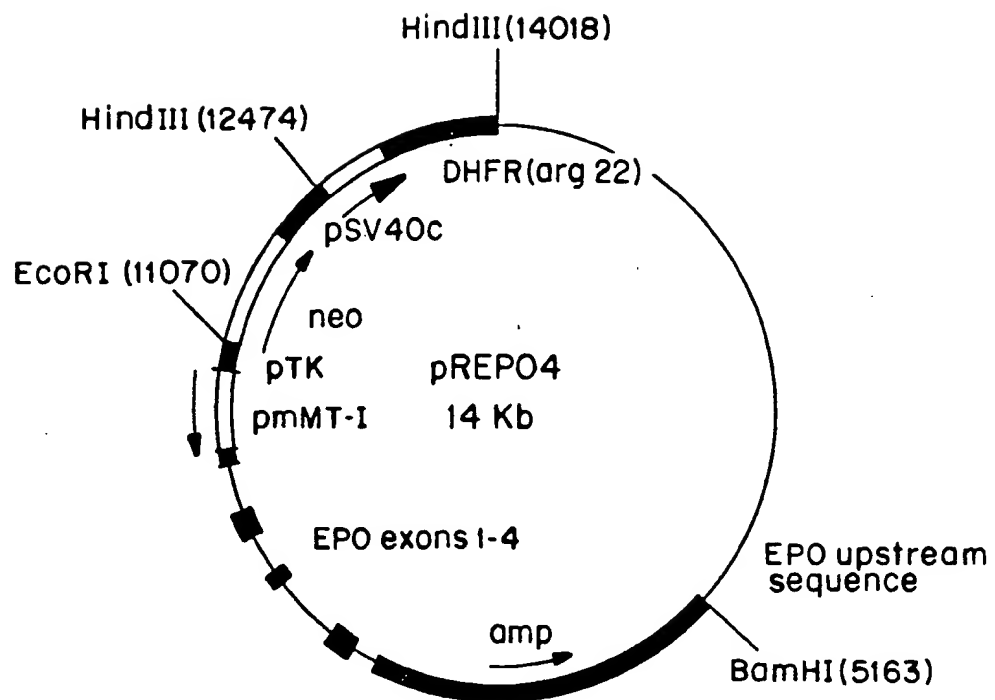


FIG. 8

-55-

resistant colonies were isolated from transfected secondary fibroblasts at a frequency of 652 G418 resistant colonies per 1.5×10^6 treated cells (1 per 2299 treated cells). Approximately 59% of these colonies express hGH.

5 EXAMPLE 2. CONSTRUCTION OF TARGETING PLASMIDS WHICH RESULT IN CHIMERIC TRANSCRIPTION UNITS IN WHICH HUMAN GROWTH HORMONE AND ERYTHROPOIETIN SEQUENCES ARE FUSED

10 The following serves to illustrate two further embodiments of the present invention, in which the normal regulatory sequences upstream of the human EPO gene are altered to allow expression of hEPO in primary or secondary fibroblast strains which do not express EPO in detectable quantities in their untransfected state as obtained.

15 In these embodiments, the products of the targeting events are chimeric transcription units in which the first exon of the human growth hormone gene is positioned upstream of EPO exons 2-5. The product of transcription, splicing and translation is a protein in which amino acids 1-4 of the

20 hEPO signal peptide are replaced with amino acid residues 1-3 of hGH. The two embodiments differ with respect to both the relative positions of the foreign regulatory sequences that are inserted and the specific pattern of splicing that needs to occur to produce the final, processed transcript.

25

Plasmid pXEPO-10 is designed to replace exon 1 of hEPO with exon 1 of hGH by gene targeting to the endogenous hEPO gene on human chromosome 7. Plasmid pXEPO-10 is constructed as follows. First, the intermediate plasmid

30 pT163 is constructed by inserting the 6 kb HindIII-BamHI fragment (see Example 1f) lying upstream of the hEPO coding region into HindIII-BamHI digested pBluescriptII SK+ (Stratagene, LaJolla, CA). The product of this ligation

-56-

tion is digested with XhoI and HindIII and ligated to the 1.1 kb HindIII-XhoI fragment from pMCIneoPolyA [Thomas, K. R. and Capecchi, M. R. Cell 51: 503-512 (1987) available from Strategene, LaJolla, CA] to create pT163. Oligo-

5 nucleotides 13.1 - 13.4 are utilized in polymerase chain reactions to generate a fusion fragment in which the mouse metallothionein 1 (mMT-I) promoter - hGH exon 1 sequences are additionally fused to hEPO intron 1 sequences. First, oligonucleotides 13.1 and 13.2 are used to amplify the

10 approximately 0.73 kb mMT-I promoter - hGH exon 1 fragment from pXGH5 (Figure 1). Next, oligonucleotides 13.3 and 13.4 are used to amplify the approximately 0.57 kb fragment comprised predominantly of hEPO intron 1 from human genomic DNA. Finally, the two amplified fragments are

15 mixed and further amplified with oligonucleotides 13.1 and 13.4 to generate the final fusion fragment (fusion fragment 3) flanked by a SalI site at the 5' side of the mMT-I moiety and an XhoI site at the 3' side of the hEPO intron 1 sequence. Fusion fragment 3 is digested with XhoI and

20 SalI and ligated to XhoI digested pT163. The ligation mixture is transformed into E. coli and a clone containing a single insert of fusion fragment 3 in which the XhoI site is regenerated at the 3' side of hEPO intron 1 sequences is identified and designated pXEPO-10.

- 25 13.1 5' AAAAGTCGAC GGTACCTTGG TTTTAAAAC CAGCCTGGAG
SalI KpnI
(SEQ ID NO 5)
- 13.2 5' CCTAGCGGCA ATGGCTACAG GTGAGTACTC GCGGGCTGGG CG
(SEQ ID NO 6)
- 30 13.3 5' CGCCCAGCCC GCGAGTACTC ACCTGTAGCC ATTGCCGCTA GG
(SEQ ID NO 7)
- 13.4 5' TTTTCTCGAG CTAGAACAGA TAGCCAGGCT GAGAG
XhoI
(SEQ ID NO 8)

-57-

The non-boldface region of oligo 13.1 is identical to the mMT-I promoter, with the natural KpnI site as its 5' boundary. The boldface type denotes a SalI site tail to convert the 5' boundary to a SalI site. The boldface region of oligos 13.2 and 13.3 denote hGH sequences, while the non-boldface regions are intron 1 sequences from the hEPO gene. The non-boldface region of oligo 13.4 is identical to the last 25 bases of hEPO intron 1. The boldface region includes an XhoI site tail to convert the 3' boundary of the amplified fragment to an XhoI site.

Plasmid pXEPO-11 is designed to place, by gene targeting, the mMT-I promoter and exon 1 of hGH upstream of the hEPO structural gene and promoter region at the endogenous hEPO locus on human chromosome 7. Plasmid pXEPO-11 is constructed as follows. Oligonucleotides 13.1 and 13.5 - 13.7 are utilized in polymerase chain reactions to generate a fusion fragment in which the mouse metallothionein I (mMT-I) promoter - hGH exon 1 sequences are additionally fused to hEPO sequences from -1 to -630 relative to the hEPO coding region. First, oligonucleotides 13.1 and 13.5 are used to amplify the approximately 0.73 kb mMT-I promoter - hGH exon 1 fragment from pXGH5 (Figure 1). Next, oligonucleotides 13.6 and 13.7 are used to amplify, from human genomic DNA, the approximately 0.62 kb fragment comprised predominantly of hEPO sequences from -1 to -620 relative to the hEPO coding region. Both oligos 13.5 and 13.6 contain a 10 bp linker sequence located at the hGH intron 1 - hEPO promoter region, which corresponds to the natural hEPO intron 1 splice donor

-58-

site. Finally, the two amplified fragments are mixed and further amplified with oligonucleotides 13.1 and 13.7 to generate the final fusion fragment (fusion fragment 6) flanked by a SalI site at the 5' side of the mMT-I moiety and an XhoI site at the 3' side of the hEPO promoter region. Fusion fragment 6 is digested with XhoI and SalI and ligated to XhoI digested pT163. The ligation mixture is transformed into *E. coli* and a clone containing a single insert of fusion fragment 6 in which the XhoI site is regenerated at the 3' side of hEPO promoter sequences is identified and designated pXEPO-11.

13.5 5' CCTAGCGGCA ATGGCTACAG GTGAGTACTC AAGCTTCTGG
HindIII
GCTTCCAGAC CCAG (SEQ ID NO 9)

15 13.6 5' CTGGGTCTGG AAGCCAGAA GCTTGAGTAC TCACCTGTAG
HindIII
CCATTGCCGC TAGG (SEQ ID NO 10)

13.7 5' TTTTCTCGAG CTCCGCGCCT GGCCGGGGTC CCTC
XhoI
(SEQ ID NO 11)

The boldface regions of oligos 13.5 and 13.6 denote hGH sequences. The italicized regions correspond to the first 10 base pairs of hEPO intron 1. The remainder of the oligos correspond to hEPO sequences from -620 to -597 relative to the hEPO coding region. The non-boldface region of oligo 13.7 is identical to bases -1 to -24 relative to the hEPO coding region. The boldface region includes an XhoI site tail

-59-

to convert the 3' boundary of the amplified fragment to an XhoI site.

Plasmid pXEPO-10 can be used for gene targeting by digestion with BamHI and XhoI to release the 7.3 kb fragment containing the mMT-I/hGH fusion flanked on both sides by hEPO sequences. This fragment (targeting fragment 1) contains no hEPO coding sequences, having only sequences lying between -620 and approximately -6620 upstream of the hEPO coding region and hEPO intron 1 sequences to direct targeting to the human EPO locus. Targeting fragment 1 is transfected into primary or secondary human skin fibroblasts using conditions similar to those described in Example 1c. G418-resistant colonies are picked into individual wells of 96-well plates and screened for EPO expression by an ELISA assay (R&D Systems, Minneapolis MN). Cells in which the transfecting DNA integrates randomly into the human genome cannot produce EPO. Cells in which the transfecting DNA has undergone homologous recombination with the endogenous hEPO intron 1 and hEPO upstream sequences contain a chimeric gene in which the mMT-I promoter and non-transcribed sequences and the hGH 5' untranslated sequences and hGH exon 1 replace the normal hEPO promoter and hEPO exon 1 (see Figure 5). Non-hEPO sequences in targeting fragment 1 are joined to hEPO sequences down-stream of hEPO intron 1. The replacement of the normal hEPO regulatory region with the mMT-I promoter will activate the EPO gene in fibroblasts, which do not normally express EPO. The replacement of hEPO exon 1 with hGH exon 1 results in a protein in which the first 4 amino acids of the hEPO signal peptide are replaced with amino acids 1-3 of hGH, creating a functional, chimeric signal peptide which is removed by post-translation pro-

-60-

cessing from the mature protein and is secreted from the expressing cells.

Plasmid pXEPO-11 can be used for gene targeting by digestion with BamHI and XhoI to release the 7.4 kb fragment containing the mMT-I/hGH fusion flanked on both sides by hEPO sequences. This fragment (targeting fragment 2) contains no hEPO coding sequences, having only sequences lying between -1 and approximately -6620 upstream of the hEPO coding region to direct targeting to the human EPO locus. Targeting fragment 2 is transfected into primary or secondary human skin fibroblasts using conditions similar to those described in Example 1g. G418-resistant colonies are picked into individual wells of 96-well plates and screened for EPO expression by an ELISA assay (R&D Systems, Minneapolis, MN). Cells in which the transfecting DNA integrates randomly into the human genome cannot produce EPO. Cells in which the transfecting DNA has undergone homologous recombination with the endogenous hEPO promoter and upstream sequences contain a chimeric gene in which the mMT-I promoter and non-transcribed sequences, hGH 5' untranslated sequences and hGH exon 1, and a 10 base pair linker comprised of the first 10 bases of hEPO intron 1 are inserted at the HindIII site lying at position -620 relative to the hEPO coding region (see Figure 6). The localization of the mMT-I promoter upstream of the normally silent hEPO promoter will direct the synthesis, in primary or secondary skin fibroblasts, of a message reading (5' to 3') non-translated metallothionein and hGH sequences, hGH exon 1, 10 bases of DNA identical to the first 10 base pairs of hEPO intron 1, and the normal hEPO promoter and hEPO exon 1 (-620 to +13 relative to the EPO coding sequence). The 10 base pair linker sequence from hEPO intron 1 acts as a splice donor site to fuse hGH exon 1 to the next downstream splice

-61-

acceptor site, that lying immediately upstream of hEPO exon 2. Processing of the resulting transcript will therefore splice out the hEPO promoter, exon 1, and intron 1 sequences. The replacement of hEPO exon 1 with hGH exon 1 results in a protein in which the first 4 amino acids of the hEPO signal peptide are replaced with amino acids 1-3 of hGH, creating a functional, chimeric signal peptide which is removed by post-translation processing from the mature protein and is secreted from the expressing cells.

5 1 results in a protein in which the first 4 amino acids of the hEPO signal peptide are replaced with amino acids 1-3 of hGH, creating a functional, chimeric signal peptide which is removed by post-translation processing from the mature protein and is secreted from the expressing cells.

10 A series of constructs related to pXEPO-10 and pXEPO-11 can be constructed, using known methods. In these constructs, the relative positions of the mMT-I promoter and hGH sequences, as well as the position at which the mMT-I/hGH sequences are inserted into hEPO upstream sequences, are varied to create alternative chimeric transcription units that facilitate gene targeting, result in more efficient expression of the fusion transcripts, or have other desirable properties. Such constructs will give similar results, such that an hGH-hEPO fusion gene is placed under the control of an exogenous promoter by gene targeting to the normal hEPO locus. For example, the 6 kb HindIII-BamHI fragment upstream of the hEPO gene (See Example 1f) has numerous restriction enzyme recognition sequences that can be utilized as sites for insertion of the neo gene and the mMT-I promoter/hGH fusion fragment. One such site, a BglII site lying approximately 1.3 kb upstream of the HindIII site, is unique in this region and can be used for insertion of one or more selectable markers and a regulatory region derived from another gene that will serve to activate EPO expression in primary, secondary, or immortalized human cells.

30 will serve to activate EPO expression in primary, secondary, or immortalized human cells.

First, the intermediate plasmid pT164 is constructed by inserting the 6 kb HindIII-BamHI fragment (Example 1f) lying upstream of the hEPO coding region into HindIII-

-62-

BamHI digested pBluescriptII SK+ (Stratagene, LaJolla, CA). Plasmid pMC1neoPolyA [Thomas, K.R. and Capecchi, M.R. Cell 51:503-512 (1987); available from Stratagene, LaJolla, CA] is digested with BamHI and XhoI, made blunt-ended by treatment with the Klenow fragment of E. coli DNA polymerase, and the resulting 1.1 kb fragment is purified. pT164 is digested with BglII and made blunt-ended by treatment with the Klenow fragment of E. coli DNA polymerase. The two preceding blunt-ended fragments are ligated together and transformed into competent E. coli. Clones with a single insert of the 1.1 kb neo fragment are isolated and analyzed by restriction enzyme analysis to identify those in which the BglII site recreated by the fusion of the blunt XhoI and BglII sites is localized 1.3 kb away from the unique HindIII site present in plasmid pT164. The resulting plasmid, pT165, can now be cleaved at the unique BglII site flanking the 5' side of the neo transcription unit.

Oligonucleotides 13.8 and 13.9 are utilized in polymerase chain reactions to generate a fragment in which the mouse metallothionein I (mMT-I) promoter - hGH exon 1 sequences are additionally fused to a 10 base pair fragment comprising a splice donor site. The splice donor site chosen corresponds to the natural hEPO intron 1 splice donor site, although a larger number of splice donor sites or consensus splice donor sites can be used. The oligonucleotides (13.8 and 13.9) are used to amplify the approximately 0.73 kb mMT-I promoter - hGH exon 1 fragment from pXGH5 (Figure 1). The amplified fragment (fragment 7) is digested with BglII and ligated to BglII digested pT165. The ligation mixture is transformed into E. coli and a clone, containing a single insert of fragment 7 in which the KpnI site in the mMT-I promoter is adjacent to the 5' end of the neo gene and the mMT-I

-63-

promoter is oriented such that transcription is directed towards the unique HindIII site, is identified and designated pXEPO-12.

13.8 5' AAAAAGATCT GGTACCTTGG TTTTAAAAC CAGCCTGGAG

5 BglIII KpnI
(SEQ ID NO 12)

10 The non-boldface region of oligo 13.8 is identical to the mMT-I promoter, with the natural KpnI site as its 5' boundary. The boldface type denotes a BglIII site tail to convert the 5' boundary to a BglIII site.

13.9 5' TTTTAGATCT GAGTACTCAC CTGTAGCCAT TGCCGCTAGG

BglIII
(SEQ ID NO 13)

15 The boldface region of oligos 13.9 denote hGH sequences. The italicized region corresponds to the first 10 base pairs of hEPO intron 1. The underlined BglIII site is added for plasmid construction purposes.

20 Plasmid pXEPO-12 can be used for gene targeting by digestion with BamHI and HindIII to release the 7.9 kb fragment containing the neo gene and the mMT-I/hGH fusion flanked on both sides by hEPO sequences. This fragment (targeting fragment 3) contains no hEPO coding sequences,
25 having only sequences lying between approximately -620 and approximately -6620 upstream of the hEPO coding region to direct targeting upstream of the human EPO locus. Targeting fragment 3 is transfected into primary, secondary, or immortalized human skin fibroblasts using conditions
30 similar to those described in Examples 1b and 1c. G418-

-64-

resistant colonies are picked into individual wells of 96-well plates and screened for EPO expression by an ELISA assay (R&D Systems, Minneapolis MN). Cells in which the transfecting DNA integrates randomly into the human genome cannot produce EPO. Cells in which the transfecting DNA has undergone homologous recombination with the endogenous hEPO promoter and upstream sequences contain a chimeric gene in which the mMT-I promoter and non-transcribed sequences, hGH 5' untranslated sequences, and hGH exon 1, and a 10 base pair linker comprised of the first 10 bases of hEPO intron 1 are inserted at the BglIII site lying at position approximately -1920 relative to the hEPO coding region. The localization of the mMT-I promoter upstream of the normally silent hEPO promoter will direct the synthesis, in primary, secondary, or immortalized human fibroblasts (or other human cells), of a message reading: (5' to 3') nontranslated metallothionein and hGH sequences, hGH exon 1, 10 bases of DNA identical to the first 10 base pairs of hEPO intron 1, and hEPO upstream region and hEPO exon 1 (from approximately -1920 to +13 relative to the EPO coding sequence). The 10 base pair linker sequence from hEPO intron 1 acts as a splice donor site to fuse hGH exon 1 to a downstream splice acceptor site, that lying immediately upstream of hEPO exon 2. Processing of the resulting transcript will therefore splice out the hEPO upstream sequences, promoter region, exon 1, and intron 1 sequences. When using pXEPO-10, -11 and -12, post-transcriptional processing of the message can be improved by using in vitro mutagenesis to eliminate splice acceptor sites lying in hEPO upstream sequences between the mMT-I promoter and hEPO exon 1, which reduce level of productive splicing events needed create the desired message. The replacement of hEPO exon 1 with hGH exon 1 results in a protein in which the first 4 amino acids of

-65-

the hEPO signal peptide are replaced with amino acids 1-3 of hGH, creating a functional, chimeric signal peptide which is removed by post-translation processing from the mature protein and is secreted from the expressing cells.

5 EXAMPLE 3. TARGETED MODIFICATION OF SEQUENCES UPSTREAM
 AND AMPLIFICATION OF THE TARGETED GENE

Human cells in which the EPO gene has been activated by the methods previously described can be induced to amplify the neo/mMT-1/EPO transcription unit if the targeting plasmid contains a marker gene that can confer resistance to a high level of a cytotoxic agent by the phenomenon of gene amplification. Selectable marker genes such as dihydrofolate reductase (dhfr, selective agent is methotrexate), the multifunctional CAD gene [encoding carbamyl phosphate synthase, aspartate transcarbamylase, and dihydro-orotase; selective agent is N-(phosphonoacetyl)-L-aspartate (PALA)], and adenosine deaminase (ada; selective agent is an adenine nucleoside), have been documented, among other genes, to be amplifiable in immortalized human cell lines (Wright, J.A. et al. Proc. Natl. Acad. Sci. USA 87:1791-1795 (1990)). In these studies, gene amplification has been documented to occur in a number of immortalized human cell lines. HT1080, HeLa, MCF-7 breast cancer cells, K-562 leukemia cells, KB carcinoma cells, or 2780AD ovarian carcinoma cells all display amplification under appropriate selection conditions.

Plasmids pXEPO-10 and pXEPO-11 can be modified by the insertion of a normal or mutant dhfr gene into the unique HindIII sites of these plasmids. After transfection of HT1080 cells with the appropriate DNA, selection for G418-resistance (conferred by the neo gene), and identification

-66-

of cells in which the hEPO gene has been activated by gene targeting of the neo, dhfr, and mMT-1 sequences to the correct position upstream of the hEPO gene, these cells can be exposed to stepwise selection in methotrexate (MTX) in order to select for amplification of dhfr and co-amplification of the linked neo, mMT-1, and hEPO sequences (Kaufman, R.J. Technique 2:221-236 (1990)). A stepwise selection scheme in which cells are first exposed to low levels of MTX (0.01 to 0.08 μ M), followed by successive exposure to incremental increases in MTX concentrations up to 250 μ M MTX or higher is employed. Linear incremental steps of 0.04 to 0.08 μ M MTX and successive 2-fold increases in MTX concentration will be effective in selecting for amplified transfected cell lines, although a variety of relatively shallow increments will also be effective. Amplification is monitored by increases in dhfr gene copy number and confirmed by measuring in vitro hEPO expression. By this strategy, substantial overexpression of hEPO can be attained by targeted modification of sequences lying completely outside of the hEPO coding region.

Constructs similar to those described (Examples 1i and 1k) to activate hGH expression in human cells can also be further modified to include the dhfr gene for the purpose of obtaining cells that overexpress the hGH gene by gene targeting to non-coding sequences and subsequent amplification.

EXAMPLE 4. TARGETING AND ACTIVATION OF THE HUMAN EPO LOCUS IN AN IMMORTALIZED HUMAN FIBROBLAST LINE

The targeting construct pXEPO-13 was made to test the hypothesis that the endogenous hEPO gene could be activated in a human fibroblast cell. First, plasmid pT22.1 was constructed, containing 63 bp of genomic hEPO sequence

-67-

upstream of the first codon of the hEPO gene fused to the mouse metallothionein-1 promoter (mMT-I). Oligonucleotides 22.1 to 22.4 were used in PCR to fuse mMT-I and hEPO sequences. The properties of these primers are as follows:

5 22.1 is a 21 base oligonucleotide homologous to a segment of the mMT-I promoter beginning 28 bp upstream of the mMT-I KpnI site; 22.2 and 22.3 are 58 nucleotide complementary primers which define the fusion of hEPO and mMT-I sequences such that the fusion contains 28 bp of

10 hEPO sequence beginning 35 bases upstream of the first codon of the hEPO gene, and mMT-I sequences beginning at base 29 of oligonucleotide 22.2, comprising the natural BglIII site of mMT-I and extending 30 bases into mMT-I sequence; 22.4 is 21 nucleotides in length and is homologous to hEPO sequences beginning 725 bp downstream of the

15 first codon of the hEPO gene. These primers were used to amplify a 1.4 kb DNA fragment comprising a fusion of mMT-I and hEPO sequences as described above. The resulting fragment was digested with KpnI (the PCR fragment contained two KpnI sites: a single natural KpnI site in the

20 mMT-I promoter region and a single natural KpnI site in the hEPO sequence), and purified. The plasmid pXEPO1 (Figure 3) was also digested with KpnI, releasing a 1.4 kb fragment and a 6.4 kb fragment. The 6.4 kb fragment was

25 purified and ligated to the 1.4 kb KpnI PCR fusion fragment. The resulting construct was called pT22.1. A second intermediate, pT22.2, was constructed by ligating the approximately 6 kb HindIII-BamHI fragment lying upstream of the hEPO structural gene (see Example 1f) to

30 BamHI and HindIII digested pBSISK+ (Stratagene, LaJolla, CA). A third intermediate, pT22.3, was constructed by first excising a 1.1 kb XhoI/BamHI fragment from pMCINEOpolyA (Stratagene, LaJolla, CA) containing the neomycin phosphotransferase gene. The fragment was then

-68-

made blunt-ended with the Klenow fragment of DNA polymerase I (New England Biolabs). This fragment was then ligated to the HincII site of pBSIISK+ (similarly made blunt with DNA polymerase I) to produce pT22.3. A fourth intermediate, pT22.4, was made by purifying a 1.1 kb XhoI/HindIII fragment comprising the neo gene from pT22.3 and ligating this fragment to XhoI and HindIII digested pT22.2. pT22.4 thus contains the neo gene adjacent to the HindIII side of the BamHI-HindIII upstream hEPO fragment.

Finally, pXEPO-13 was generated by first excising a 2.0 kb EcoRI/AccI fragment from pT22.1. The EcoRI site of this fragment defines the 5' boundary of the mMT-I promoter, while the AccI site of this fragment lies within hEPO exon 5. Thus, the AccI/EcoRI fragment contains a nearly complete hEPO expression unit, missing only a part of exon 5 and the natural polyadenylation site. This 2.0 kb EcoRI/AccI fragment was purified, made blunt-ended by treatment with the Klenow fragment of DNA polymerase I, and ligated to XhoI digested, blunt-ended, pT22.4.

HT1080 cells were transfected with PvuI-BamHI digested pXEPO-13. pXEPO-13 digested in this way generates three fragments; a 1 kb vector fragment including a portion of the amp gene, a 1.7 kb fragment of remaining vector sequences and an approximately 10 kb fragment containing hEPO, neo and mMT-I sequences. This approximately 10 kb BamHI/PvuI fragment contained the following sequences in order from the BamHI site: an approximately 6.0 kb of upstream hEPO genomic sequence, the 1.1 kb neo transcription unit, the 0.7 kb mMT-I promoter and the 2.0 kb fragment containing hEPO coding sequence truncated within exon 5. 45µg of pXEPO-13 digested in this way was used in an electroporation of 12 million cells (electroporation conditions were described in Example 1b). This electroporation was repeated a total of eight times,

-69-

resulting in electroporation of a total of 96 million cells. Cells were mixed with media to provide a cell density of 1 million cells per ml and 1 ml aliquots were dispensed into a total of 96, 150mm tissue culture plates (Falcon) each containing a minimum of 35 ml of DMEM/15% calf serum. The following day, the media was aspirated and replaced with fresh medium containing 0.8 mg/ml G418 (Gibco). After 10 days of incubation, the media of each plate was sampled for hEPO by ELISA analysis (R & D Systems). Six of the 96 plates contained at least 10 mU/ml hEPO. One of these plates, number 18, was selected for purification of hEPO expressing colonies. Each of the 96, 150 mm plates contained approximately 600 G418 resistant colonies (an estimated total of 57,600 G418 resistant colonies on all 96 plates). The approximately 600 colonies on plate number 18 were trypsinized and replated at 50 cells/ml into 364 well plates (Sterilin). After one week of incubation, single colonies were visible at approximately 10 colonies per large well of the 364 well plates (these plates are comprised of 16 small wells within each of the 24 large wells). Each well was screened for hEPO expression at this time. Two of the large wells contained media with at least 20 mU/ml hEPO. Well number A2 was found to contain 15 colonies distributed among the 16 small wells. The contents of each of these small wells were trypsinized and transferred to 16 individual wells of a 96 well plate. Following 7 days of incubation the media from each of these wells was sampled for hEPO ELISA analysis. Only a single well, well number 10, contained hEPO. This cell strain was designated HT165-18A2-10 and was expanded in culture for quantitative hEPO analysis, RNA isolation and DNA isolation. Quantitative measurement of hEPO production resulted in a value of 2,500 milliunits/million cells/24 hours.

-70-

A 0.2 kb DNA probe extending from the AccI site in hEPO exon 5 to the BglII site in the 3' untranslated region was used to probe RNA isolated from HT165-18A2-10 cells. The targeting construct, pXEPO-13, truncated at the AccI site in exon 5 does not contain these AccI/BglII sequences and, therefore, is diagnostic for targeting at the hEPO locus. Only cell strains that have recombined in a homologous manner with natural hEPO sequences would produce an hEPO mRNA containing sequence homologous to the AccI/BglII sequences. HT165-18A2-10 was found to express an mRNA of the predicted size hybridizing with the 32-P labeled AccI/BglII hEPO probe on Northern blots. Restriction enzyme and Southern blot analysis confirmed that the neo gene and mMT-I promoter were targeted to one of the two hEPO alleles in HT165-18A2-10 cells.

These results demonstrate that homologous recombination can be used to target a regulatory region to a gene that is normally silent in human fibroblasts, resulting in the functional activation of that gene.

- 22.1 5' CACCTAAAAT GATCTCTCTG G (SEQ ID NO 14)
- 22.2 5' CGCGCCGGGT GACCACACCG GGGGCCCTAG ATCTGGTGAA
GCTGGAGCTA CGGAGTAA (SEQ ID NO 15)
- 22.3 5' TTA CTCCGTA GCTCCAGCTT CACCAGATCT AGGGCCCCCG
GTGTGGTCAC CCGGCGCG (SEQ ID NO 16)
- 22.4 5' GTCTCACCGT GATATTCTCG G (SEQ ID NO 17)

EXAMPLE 5. PRODUCTION OF INTRONLESS GENES

Gene targeting can also be used to produce a processed gene, devoid of introns, for transfer into yeast or bacteria for gene expression and in vitro protein produc-

-71-

tion. For example, hGH can be produced in yeast by the approach described below.

Two separate targeting constructs are generated. Targeting construct 1 (TC1) includes a retroviral LTR sequence, for example the LTR from the Moloney Murine Leukemia Virus (MoMLV), a marker for selection in human cells (e.g., the neo gene from Tn5), a marker for selection in yeast (e.g., the yeast URA3 gene), a regulatory region capable of directing gene expression in yeast (e.g., the GAL4 promoter), and optionally, a sequence that, when fused to the hGH gene, will allow secretion of hGH from yeast cells (leader sequence). The vector can also include a DNA sequence that permits retroviral packaging in human cells. The construct is organized such that the above sequences are flanked, on both sides, by hGH genomic sequences which, upon homologous recombination with genomic hGH gene N sequences, will integrate the exogenous sequences in TC1 immediately upstream of hGH gene N codon 1 (corresponding to amino acid position 1 in the mature, processed protein). The order of DNA sequences upon integration is: hGH upstream and regulatory sequences, neo gene, LTR, URA3 gene, GAL4 promoter, yeast leader sequence, hGH sequences including and downstream of amino acid 1 of the mature protein. Targeting Construct 2 (TC2) includes sequences sufficient for plasmid replication in yeast (e.g., 2-micron circle or ARS sequences), a yeast transcriptional termination sequence, a viral LTR, and a marker gene for selection in human cells (e.g., the bacterial gpt gene). The construct is organized such that the above sequences are flanked on both sides by hGH genomic sequences which, upon homologous recombination with genomic hGH gene N sequences, will integrate the exogenous sequences in TC2 immediately downstream of the hGH gene N stop codon. The order of DNA sequences upon

-72-

integration is: hGH exon 5 sequences, yeast transcription termination sequences, yeast plasmid replication sequences, LTR, gpt gene, hGH 3' non-translated sequences.

Linear fragments derived from TC1 and TC2 are sequentially targeted to their respective positions flanking the hGH gene. After superinfection of these cells with helper retrovirus, LTR directed transcription through this region will result in an RNA with LTR sequences on both ends. Splicing of this RNA will generate a molecule in which the normal hGH introns are removed. Reverse transcription of the processed transcript will result in the accumulation of double-stranded DNA copies of the processed hGH fusion gene. DNA is isolated from the doubly-targeted, retrovirally-infected cells, and digested with an enzyme that cleaves the transcription unit once within the LTR. The digested material is ligated under conditions that promote circularization, introduced into yeast cells, and the cells are subsequently exposed to selection for the URA3 gene. Only cells which have taken up the URA3 gene (linked to the sequences introduced by TC1 and TC2 and the processed hGH gene) can grow. These cells contain a plasmid which will express the hGH protein upon galactose induction and secrete the hGH protein from cells by virtue of the fused yeast leader peptide sequence which is cleaved away upon secretion to produce the mature, biologically active, hGH molecule.

Expression in bacterial cells is accomplished by simply replacing, in TC1 and TC2, the ampicillin-resistance gene from pBR322 for the yeast URA3 gene, the tac promoter (deBoer *et al.*, Proc. Natl. Acad. Sci. 80:21-25 (1983)) for the yeast GAL4 promoter, a bacterial leader sequence for the yeast leader sequence, the pBR322 origin of replication for the 2-micron circle or ARS sequence, and a bacterial transcriptional termination (e.g., trpA

-73-

transcription terminator; Christie, G.E. *et al.*, Proc. Natl. Acad. Sci. 78:4180-4184 (1981)) sequence for the yeast transcriptional termination sequence. Similarly, hEPO can be expressed in yeast and bacteria by simple
5 replacing the hGH targeting sequences with hEPO targeting sequences, such that the yeast or bacterial leader sequence is positioned immediately upstream of hEPO codon 1 (corresponding to amino acid position 1 in the mature processed protein).

10 EXAMPLE 6. ACTIVATION AND AMPLIFICATION OF THE EPO GENE
IN AN IMMORTALIZED HUMAN CELL LINE

Incorporation of a dhfr expression unit into the unique HindIII site of pXEPO-13 (see Example 4) results in a new targeting vector capable of dual selection and
15 selection of cells in which the dhfr gene is amplified. The single HindIII site in pXEPO-13 defines the junction of the neo gene and genomic sequence naturally residing upstream of the human EPO gene. Placement of a dhfr gene at this site provides a construct with the neo and dhfr
20 genes surrounded by DNA sequence derived from the natural EPO locus. Like pXEPO-13, is useful to target to the EPO locus by homologous recombination. Such a construct designated pREPO4, is represented in Figure 8. The plasmid includes exons 1-4 and part of exon 5 of the human EPO
25 gene, as well as the HindIII-BamHI fragment lying upstream of the hEPO coding region. pSVe, pTK and pmMT-I correspond to the promoters from the SV40 early region, the Herpes Simplex Virus (HSV) thymidine kinase (TK) gene and the mouse metallothionein-I gene. It was produced as
30 follows: HindIII-digested pXEPO-13 was purified and made blunt with the Klenow fragment of DNA polymerase I. To obtain a dhfr expression unit, the plasmid construct pF8CIS9080 (Eaton *et al.*, Biochemistry 25:8343-8347

-74-

(1986)) was digested with EcoRI and SalI. A 2 Kb fragment containing the dhfr expression unit was purified from this digest and made blunt with Klenow fragment of DNA polymerase I. This dhfr-containing fragment was then ligated to the blunted HindIII site of pXEPO-13. An aliquot of this ligation was transformed into E. coli and plated on ampicillin selection plates. Following an overnight incubation at 37°C, individual bacterial colonies were observed, picked and grown. Miniplasmid preparations were made from these cultures and the resulting DNA was then subjected to restriction enzyme digestion with the enzymes BglI+HindIII, and SfiI in order to determine the orientation of the inserted dhfr fragments. Plasmid DNA from one of these preparations was found to contain such a 2 Kb insertion of the dhfr fragment. The transcription orientation of the dhfr expression unit in this plasmid was found to be opposite that of the adjacent neo gene. This is the construct designated pREPO4.

Plasmid pREPO4 was used to amplify the hEPO locus in cells subsequent to activation of the endogenous hEPO gene by homologous recombination. Gene activation with this construct allows selection for increased DHFR expression by the use of the drug methotrexate (MTX). Typically, increased DHFR expression would occur by an increase in copy number through DNA amplification. The net result would be co-amplification of the activated EPO gene along with dhfr sequences. Co-amplification of the activated EPO locus should result in increased EPO expression.

Targeting experiments were performed in HT1080 cells with pREPO4. hEPO expressing line HTREPO-52 was isolated. This line was analyzed quantitatively for EPO production and by Southern and Northern blot. Both strains were found to be targeted with a single copy of dhfr/neo/mMT-1 sequences. Expression levels obtained under 0.8 mg/ml

-75-

- G418 selection were approximately 1300 mU/million cells/day. Because the targeted EPO locus contained a dhfr expression unit, it was possible to select for increased expression of DHFR with the antifolate drug, MTX.
- 5 These strains were therefore subjected to stepwise selection in 0.02, 0.05, 0.1, 0.2 and 0.4 μ M MTX. Results of initial selection of mass cultures of these strains are shown in Table 4 and Figure 7.

TABLE 4

<u>Cell Line</u>	<u>MTX (μM)</u>	<u>mU/ Million Cells/ 24h</u>
52C20-5-0	0	1368
52C20-5-.01	0.01	1744
52C20-5-.02	0.02	11643
52C20-5-0.05	0.05	24449
52-3-5-0.10	0.1	37019
52-3-2-0.20	0.2	67867
52-3-2-0.4B	0.4	99919

- Selection with elevated levels of MTX was successful in increasing EPO expression in line HTREPO-52, with a 70-fold increase in EPO production seen in the cell line resistant to 0.4 μ M MTX. Confirmation of amplification of
- 5 the EPO locus was accomplished by Southern blot analysis in MTX-resistant cell lines, which revealed an approximately 10-fold increase in the copy number of the activated hEPO locus relative to the parental (untargeted) hEPO allele.

-76-

CLAIMS

1. A method of activating expression of and amplifying an endogenous gene in genomic DNA of a vertebrate cell which is not expressed in the cell as obtained or is not expressed at significant levels in the cell as obtained, comprising the steps of:
- 5 a) transfecting cells with DNA sequences comprising:
- 1) exogenous DNA selected from the group consisting of:
- 10 a) DNA sequences which repair, alter, delete or replace a sequence present in the cell; or
- 15 b) DNA sequences which are regulatory sequences not normally functionally linked to the endogenous gene in the cell as obtained;
- 2) DNA sequences homologous with genomic DNA sequences at a preselected site in the cells; and
- 20 3) amplifiable DNA encoding a selectable marker,
- thereby producing cells containing the DNA sequences;
- b) maintaining the cells produced in (a) under conditions appropriate for homologous recombination to occur between DNA sequences homologous
- 25 with genomic DNA sequences and genomic DNA sequences,
- thereby producing homologously recombinant cells of vertebrate origin having the DNA sequences of (a)(1),
- 30 (a)(2) and (a)(3) integrated into genomic DNA and exogenous DNA of (a)(1) functionally linked to the endogenous gene; and

-77-

- 5 c) culturing the homologously recombinant cells produced in (b) under conditions which select for amplification of the amplifiable DNA encoding a selectable marker, whereby the amplifiable DNA encoding a selectable marker and the endogenous gene functionally linked to exogenous DNA of (a)(1) are coamplified, thereby producing homologously recombinant cells containing amplified DNA encoding a selectable marker and a coamplified endogenous gene functionally linked to DNA sequences of (a)(1), in which the coamplified gene is expressed.
- 10
2. A method of targeting DNA sequences into genomic DNA of a cell of vertebrate origin, comprising the steps of:
- 15
- a) providing a DNA construct comprising:
- 1) exogenous DNA encoding a product to be expressed in cells of vertebrate origin;
- 2) DNA sequences homologous with genomic DNA sequences in the cell of vertebrate origin; and
- 20
- 3) amplifiable DNA sequences encoding a marker for which selection of amplified copies of the marker can be performed;
- 25
- b) introducing into cells of vertebrate origin the DNA construct provided in (a), thereby producing cells containing the DNA construct provided in (a);
- 30
- c) maintaining cells produced in (b) under conditions appropriate for homologous recombination to occur between DNA sequences homologous with genomic DNA sequences and genomic DNA sequences,

-78-

thereby producing homologously recombinant cells of vertebrate origin having the DNA construct of (a) integrated into genomic DNA of the cells; and

- 5 d) culturing the homologously recombinant cells produced in (b) under conditions which select for amplification of the amplifiable DNA encoding a selectable marker, whereby the amplifiable DNA encoding a selectable marker and the exogenous DNA of (a)(1) are coamplified,
- 10 thereby producing homologously recombinant cells containing amplified DNA encoding a selectable marker and a coamplified exogenous gene wherein the cells are capable of expressing the coamplified exogenous gene.

- 15 3. The method of Claim 1 or Claim 2 wherein the vertebrate cell is a primary or secondary cell.
4. The method of Claim 3 wherein the primary or secondary cell is a mammalian cell.
5. The method of Claim 3 wherein the primary or secondary
- 20 cell is a human cell.
6. The method of Claim 1 or Claim 2 wherein the vertebrate cell is an immortalized cell.
7. The method of Claim 6 wherein the immortalized cell is of mammalian origin.
- 25 8. The method of Claim 6 wherein the immortalized cell is of human origin.

-79-

9. The method of any one of Claims 1-8 wherein the amplifiable DNA encoding a selectable marker encodes a selectable marker selected from the group consisting of: dihydrofolate reductase, adenosine deaminase, and CAD.
- 5
10. The method of any one of Claims 1-9 in which the DNA construct further comprises (a) an additional positive selection marker, (b) a negative selection marker or (c) an additional positive selection marker and a negative selection marker.
- 10
11. The method of Claim 10 in which the additional positive selection marker is neo.
12. The method of Claim 10 in which the positive selection marker is neo and the negative selection marker is gpt or the HSV-TK gene.
- 15
13. The method of any one of Claims 1-12 wherein the gene to be expressed encodes a product selected from the group consisting of: hormones, cytokines, antigens, antibodies, enzymes, clotting factors, transport proteins, receptors, regulatory proteins, structural proteins, transcription factors, antisense RNA and ribozymes, and proteins or nucleic acids which do not occur in nature.
- 20
14. The method of any one of Claims 1-12 wherein the endogenous gene encodes a therapeutic product selected from the group consisting of human growth hormone, human insulin, human insulinotropin, and human erythropoietin.
- 25

-80-

15. Homologously recombinant cells produced by the method of any of the preceding Claims.
- 5 16. The method of any one of Claims 1 to 8 wherein the endogenous gene or the exogenous gene to be expressed encodes a therapeutic product.
- 10 17. The method of Claim 16 in which the DNA construct further comprises (a) an additional positive selection marker, (b) a negative selection marker or (c) an additional positive selection marker and a negative selection marker.
18. The method of Claim 17 in which the additional positive selection marker is neo.
- 15 19. The method of Claim 17 in which the positive selection marker is neo and the negative selection marker is gpt or the HSV-TK gene.
- 20 20. The method of any one of Claims 16-19 wherein the gene to be expressed encodes a product selected from the group consisting of: hormones, cytokines, antigens, antibodies, enzymes, clotting factors, transport proteins, receptors, regulatory proteins, structural proteins, transcription factors, antisense RNA and ribozymes, and proteins or nucleic acids which do not occur in nature.
- 25 21. The method of Claim 20 wherein the endogenous gene or the exogenous gene encodes a therapeutic product selected from the group consisting of human growth hormone, human insulin, human insulinotropin, and human erythropoietin.

-81-

22. Homologously recombinant cells produced by the method of any of Claims 16-21.
23. Primary cells or secondary cells produced by the method of any one of Claims 1-5 and 9-14 and for use in therapy, e.g. mammalian or human therapy.
24. Use of cells produced by the method of any one of Claims 1-5 and 9-14 for the manufacture of a medicament for use in gene therapy.
25. A barrier device adapted to be introduced into a mammal and enclosing cells produced by the method of Claim 16, wherein the barrier device is made of a material which when the device is introduced into the mammal permits passage of the therapeutic product into the circulation or tissues of the mammal while preventing the vertebrate cells from growing outside of the barrier device and the rejection of the cells by mammal's immune system.
26. A method of providing an effective amount of a therapeutic product to a mammal comprising introducing into the mammal cells produced by the method of Claim 16, in which the cells are introduced into the mammal enclosed within a barrier device, wherein the barrier device is made of a material which permits passage of the therapeutic product into the circulation or tissues of the mammal and prevents the vertebrate cells from growing outside of the barrier device and prevents the rejection of the cells by the mammal's immune system.

-82-

27. A method of providing an effective amount of a therapeutic product to a mammal comprising introducing into the mammal cells produced by the method of Claim 16.
- 5 28. A method of providing an effective amount of a therapeutic product to a human comprising introducing into the human cells produced by the method of Claim 16.
- 10 29. A method for the in vitro production of a therapeutic protein comprising culturing the homologously recombinant cells Claim 9 under conditions suitable for the expression of the therapeutic protein, whereby the therapeutic protein is expressed by the cells of Claim 9.
- 15 30. A method for the in vitro production of a therapeutic protein using a microbial cell containing an intronless copy of a human gene encoding a therapeutic protein, comprising the steps of:
- a) introducing into human cells a DNA construct comprising:
- 20 1) a retroviral LTR;
- 2) a gene encoding a marker for selection in microbial cells;
- 3) regulatory sequences capable of promoting expression of the therapeutic protein in microbial cells;
- 25 4) DNA encoding a leader peptide capable of promoting secretion of the therapeutic protein from microbial cells; and
- 5) sequences sufficient to direct homologous recombination of the DNA construct with human genomic DNA sequences adjacent to and
- 30

-83-

upstream of DNA encoding the first amino acid in the therapeutic protein, thereby producing cells containing the DNA construct;

- 5 b) maintaining the cells produced in (a) under conditions appropriate for homologous recombination to occur between DNA sequences homologous with genomic DNA sequences and genomic DNA sequences, thereby producing homologously recombinant cells having the DNA construct of (a) integrated into genomic DNA adjacent to and upstream of and functionally linked to the gene encoding the therapeutic protein;
- 10 c) introducing into the homologously recombinant cells produced in (b) a DNA construct comprising:
- 15 1) DNA sequences capable of directing transcription termination in a microbial cell;
- 2) DNA sequences capable of directing DNA replication in a microbial cell;
- 20 3) a retroviral LTR;
- 4) sequences sufficient to direct homologous recombination of the DNA construct with human genomic DNA sequences downstream of the stop codon of the therapeutic protein,
- 25 thereby producing homologously recombinant cells having integrated the DNA construct of (a) into the genomic DNA, and containing the DNA construct of (c);
- 30 d) maintaining the cells produced in (c) under conditions appropriate for homologous recombination to occur between DNA sequences homologous with genomic DNA sequences and genomic DNA sequences, thereby producing homologously recombinant cells having the DNA construct of (a) inte-

-84-

- grated into genomic DNA upstream of and functionally linked to the gene encoding the therapeutic protein, and also having the DNA construct (c) integrated downstream of and functionally linked to the gene encoding the therapeutic protein;
- 5
- e) culturing the cells produced in (d) under conditions appropriate for retroviral LTR directed transcription, processing and reverse transcription of the RNA product of the DNA sequences between the LTRs introduced in (a) and (c), thereby producing a DNA sequence comprising an intronless DNA copy of the gene encoding the therapeutic protein operatively linked to the DNA sequences comprising the DNA construct described in (a) and operatively linked to the DNA sequences comprising the DNA construct described in (c), this sequence herein referred to as the intronless DNA sequence;
- 10
- 15
- 20 f) separating the intronless DNA sequence produced in (e) from the cells;
- g) introducing the DNA separated in (f) into microbial cells, thereby producing microbial cells containing the DNA produced in (e);
- 25
- h) culturing the microbial cells produced in (g) in the presence of a selective agent which selects for the selectable marker present in the DNA construct described in (a);
- 30
- i) culturing the cells produced in (h) under conditions appropriate for expression and secretion of the therapeutic protein in microbial cells, thereby producing the therapeutic protein in the microbial cells and secreting the protein therefrom; and

-85-

- j) separating the therapeutic protein produced in
(i) from the microbial cells.

31. The method of Claim 30 wherein the microbial cell is
Saccharomyces cerevisiae, or Escherichia coli.

5 32. A method of turning on expression of a gene to be
expressed which is present in a cell but is not ex-
pressed in the cell as obtained or is not expressed
at significant levels in the cell as obtained, com-
prising introducing into the cell a DNA construct
10 comprising a regulatory region under conditions
appropriate for homologous recombination, whereby the
regulatory region is inserted into or replaces all or
a portion of the regulatory region of the gene to be
expressed, and is functionally linked to the gene to
15 be expressed, thereby producing homologously recombi-
nant cells which express the gene.

33. The method of Claim 32 wherein the cell as obtained
is an immortalized cell.